DOCUMENT RESUME

ED 197 043 UD 021 187

AUTHOR Keesling, J. Ward: And Others
TITLE Compensatory Education and Confined Youth. Volume 2: A National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth.
INSTITUTION System Development Corp., Santa Monica, Calif. Studies and Evaluation Dept.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO TM-5792/011/00
PUB DATE Jun 79
CONTRACT 300-76-0093
NOTE 271p.: For related documents see UD 021 188-189 and UD 021 220.

EDRS PRICE MF01/PC11 Plus Postage.
DESCRIPTORS *Academic Achievement: *Basic Skills: Child Neglect: *Compensatory Education: Correctional Education: Correctional Institutions: Delinquency: Delinquent Rehabilitation: Educational Improvement: Institutional Characteristics: *Institutionalized Persons: Program Evaluation: Secondary Education: *Student Attitudes: *Test Results
IDENTIFIERS *Elementary Secondary Education Act Title I

ABSTRACT

This study (one portion of a larger study) was designed to determine whether Elementary and Secondary Education Act Title I programs in State institutions for neglected or delinquent youth are effective in improving students' basic skills and attitudes and to determine which components of educational programs and characteristics of institutions were related to such improvements. A sample of 40 institutions was surveyed. Data were collected using domain referenced tests of basic reading and arithmetic and standardized achievement tests. Descriptive data were collected from institutional staff members and teachers' records. Findings indicated that institutionalized Title I students performed poorly on tests of basic skills and that achievement levels and attitudes remained stable across testing periods. Data also showed that Title I students were offered less reading instructional time and more math instructional time than were non-Title I students, but low attendance rates reduced the actual amount of instruction received. The costs of Title I services were found to be higher than the costs of regular educational programs. Analyses indicated that institutions able to devote time and resources directly to programs aimed at the achievement and attitudinal needs of students were more likely to be successful in effecting changes in their students. (Author/MK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED197043

COMPENSATORY EDUCATION AND CONFINED YOUTH:

A National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth.

VOLUME 2

J. Ward Keesling
Noreen M. Webb
Judy C. Pfannenstiel

June 1979

This study was supported by and forms a part of the National Evaluation of Title I Programs for Neglected or Delinquent Youth in State Institutions. United States Office of Education Contract Number 300-76-0093

TM-5792/011/00

Education Studies and Evaluation Department
System Development Corporation
Santa Monica, California

UD021187

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Project Officer, Consultants
and Resource Persons and
Project Personnel

Project Officer

Janice Anderson

Consultants and Resource Persons

Eric C. Brookens
Thomas J. Cook
Joyce DeMuth
LaMar T. Empey
Edmund J. Gubbins
Chester W. Harris
Tina M. Jackson
Jeremy J. Kollas
John B. Lampe
Robert L. Linn
William Lobosco
Patrick O. Mancini
Lawrence D. Penny
Melvin Seeman
George L. Smith
Frances A. Tyler
Mildred R. Walton
James R. Wickman
Alan Woods

Project Co-Directors

J. Ward Keesling
Judy C. Pfannenstiel

Former Project Staff

Ted Bartell, former SDC
Project Director
Sheila Cottral
Pamela DeFusco
Noreen M. Webb

Project Secretary

M. Helen Hernandez

Field Data Collection

James Hoyt, Manager of
Field Data Collection
Rayden Acevedo
Claudia Browne
Sue Brown
Cathie Feulner Fa[illegible]ar
Mel McCane
Doug Robertson
Cathe Talavera
Jeffrey Wallace
Renee Wolff

Acknowledgements

We wish to thank several people for their contributions to this study. Dr. Ted Bartell, former SDC Project Director, developed the design of this research and constructed the sampling frame of institutions from which the sites were drawn. Dr. William Doherty emphasized the importance of domain referenced testing for this study and developed the first versions of the tests. James Hoyt ably managed the field data collection. Sheila Cottral cleaned the data on student participation and coded the domain referenced test in division.

The preparation of this manuscript required the editorial assistance of Frank Tierney and Pamela Doherty and the tireless typing of M. Helen Hernandez and Marcia Simonson.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	xvi
EXECUTIVE SUMMARY	xvii
CHAPTER 1. INTRODUCTION	1
BACKGROUND AND OBJECTIVES OF THE OVERALL STUDY	1
PHASE I OBJECTIVES	2
OVERVIEW OF PHASE I FINDINGS	3
Characteristics of the Students Served by Title I Programs	3
Content of the Title I Program	4
Title I Setting	4
Program Management	5
Title I Funding	7
Opinions Concerning the Operation and Effectiveness of Title I Programs	7
PHASE II OBJECTIVES	9
BRIEF GUIDE FOR THE READER OF VOLUME II	9
CHAPTER II. DESIGN OF PHASE II IMPACT EVALUATION	11
PRECIS OF CHAPTER II	11
OVERVIEW OF THE STUDY DESIGN	12
SAMPLING PLAN	13
FIELD OPERATIONS	18
RATES OF STUDENT PARTICIPATION IN TESTING	21
CHAPTER III. INSTRUMENTATION FOR THE PHASE II STUDY	27
PRECIS OF CHAPTER III.	27
OVERVIEW OF INSTRUMENT DEVELOPMENT	29
DESCRIPTIVE INFORMATION	30
Educational Program Characteristics	30
Institutional Characteristics	32
Student Characteristics	32

TABLE OF CONTENTS (Cont'd)

	Page
STUDENT ACHIEVEMENT	32
Standardized Achievement Measures	32
Domain Referenced Tests (DRTs)	33
The Domain Referenced Test of Arithmetic Basic Skills	35
The Domain Referenced Tests of Language Arts Basic Skills	38
Form-to-Form Difficulty and Reliability of the Domain-Referenced Test	39
STUDENT ATTITUDES	43
Development of Student Attitude Measures	43
CHAPTER IV. CHARACTERISTICS OF STUDENTS, PROGRAMS, AND INSTITUTIONS	49
PRECIS OF CHAPTER IV	49
STUDENT CHARACTERISTICS	51
DESCRIPTIONS OF INSTITUTIONS	56
CHARACTERISTICS OF EDUCATIONAL PROGRAMS	58
Residents in Programs	58
Operating Costs of Programs	60
Distribution of Teaching Assignments	65
Teaching Practice and Classroom Support	67
Comparison of Selected Teaching Practices With Previous Findings	71
Assignment of Students to Educational Programs	73
Provision of Instruction in Basic Skills	77
CHAPTER V. THE RELATIONSHIP OF THE CHARACTERISTICS OF INSTITUTIONS AND EDUCATIONAL PROGRAMS TO THE ACHIEVEMENT AND ATTITUDES OF INSTITUTIONALIZED YOUTH	83
PRECIS OF CHAPTER V	83
STANDARDIZED ACHIEVEMENT TEST RESULTS	85
Practical Achievement Test	85
Comprehensive Test of Basic Skills	85
RESULTS OF DOMAIN REFERENCED TESTS	89
ANALYSIS OF PURPOSIVELY SELECTED SITES	108

TABLE OF CONTENTS (Cont'd)

	Page
STUDENT ATTITUDES	108
Self Esteem and Locus of Control	108
Student Attitudes Toward School	120
Student Perceptions of the Educational Environment	121
THE RELATIONSHIP OF CHARACTERISTICS OF EDUCATIONAL PROGRAMS TO ACHIEVEMENT AND ATTITUDE MEASURES	121
Analytical Process	121
Results of Relational Analyses - Achievement Scores	123
Results of Relational Analyses - Attitudinal Measures	127
Self-Esteem	127
Locus of Control	139
Attitudes Toward Reading	139
Attitudes Toward Mathematics	140
Affiliation	140
School Involvement	141
Liking of School	141
Sociability	142
Teacher Supportiveness	142
Practical Orientation of Learning	143
Post-Release Expectations and Stigmatization	143
CONCLUSIONS AND RECOMMENDATIONS	144
REFERENCES	149
APPENDIX A. DEVELOPMENT OF THE DOMAIN-REFERENCED TESTS	A-1
DOMAIN IDENTIFICATION AND DEFINITION	A-1
DEVELOPMENT OF THE LANGUAGE ARTS TESTS	A-9
Domain Specification	A-9
DEVELOPMENT OF THE MATHEMATICS TESTS	A-20
Domain Specification	A-20

TABLE OF CONTENTS (Cont'd)

	Page
APPENDIX B. HIERARCHICAL STRUCTURE OF THE DOMAIN-REFERENCED TESTS . . .	B-1
APPENDIX C. RESULTS OF THE FACTOR ANALYSIS OF STUDENT ATTITUDE SCALES .	C-1
APPENDIX D. AVERAGE STUDENT BACKGROUND CHARACTERISTICS AND ATTITUDES FOR EACH OF 30 RANDOMLY SAMPLED INSTITUTIONS	D-1
APPENDIX E. DISTRIBUTION OF RESPONSES OF TITLE I AND REGULAR EDUCATION TEACHERS	E-1
APPENDIX F. DISTRIBUTION OF TEACHER RESPONSES WITHIN INSTITUTIONS . . .	F-1
APPENDIX G. MEAN PROPORTIONS CORRECT ON DOMAIN REFERENCED ACHIEVEMENT TESTS FOR EACH RANDOMLY SELECTED INSTITUTION	G-1
APPENDIX H. IS TITLE I EFFECTIVE IN IMPROVING STUDENTS' ATTITUDES AND PROFICIENCY IN BASIC SKILLS? RESULTS OF ANALYSES-OF-COVARIANCE	H-1
APPENDIX I. A COMPARISON OF THE DOMAIN REFERENCED TEST SCORES OF NEGLECTED OR DELINQUENT YOUTH WITH SCORES OF ELEVENTH GRADERS FROM THE FLORIDA STATEWIDE ASSESSMENT PROGRAM	I-1

LIST OF TABLES

		Page
Table II-1.	Content of Instruments Used in the Phase II Study	14
Table II-2.	Number of Facilities Eligible for Selection into Phase II Samples, by Strata	16
Table II-3.	Number of Facilities in Phase II Sample, by Strata	16
Table II-4.	Contents of the Procedures Manual	19
Table II-5.	Schedule of Data Collection Activities for Phase II Study	22
Table II-6.	Percentage of Students Selected for First Testing Who Were Present at Institutions for Later Testings	24
Table II-7.	Students Selected for Testing During Wave I Who Were Tested at Subsequent Site Visits	24
Table III-1.	Mean Proportion Correct on Arithmetic Subtests at Wave 1 .	37
Table III-2.	Mean Proportion Correct on Literal Comprehension Passages at Wave 1	38
Table III-3.	Differences Across Parallel Forms of the Domain-Referenced Tests at Wave 1	40
Table III-4.	Reliability of Domain-Referenced Tests	41
Table III-5.	Dimensions of Student Perceptions of the Educational Programs Within the Institution	45
Table III-6.	Student Attitude Scales	47
Table IV-1.	Distribution of Student Background Characteristics by Educational Program	52
Table IV-2.	Percent of Students in Each Category of Reason for Institutionalization and Educational Program	53
Table IV-3.	Percent of Students in Each Category of Ethnic Background and Educational Program	54
Table IV-4.	Descriptive Information on 30 Randomly Sampled State-Operated Institutions for the Neglected or Delinquent . .	57
Table IV-5.	Participation of Residents in Educational Programs	59

LIST OF TABLES (Cont'd)

		Page
Table IV-6.	Other Characteristics of Educational Programs	61
Table IV-7.	Summary of Data From Fiscal Officers at 30 Randomly Sampled Institutions	62
Table IV-8.	Summary of Data on Costs Per Student at 30 Randomly Sampled Institutions	63
Table IV-9.	Distribution of Teaching Assignments	65
Table IV-10.	Distribution of Basic Skills Assignments	66
Table IV-11.	Distribution of Mean Teacher Responses on Variables Found to Distinguish Among Institutions	69
Table IV-12.	Average Variation of Teacher Responses Within Institutions	70
Table IV-13.	Comparison of Selected Responses of Reading Teachers to Previous Findings	72
Table IV-14.	Basis for Selecting Students to Participate in Educational Programs	73
Table IV-15.	Basis for Assigning Students to Title I Programs	74
Table IV-16.	Minutes of Instruction Per Week in Basic Skills Subjects	79
Table IV-17.	Percentages of Scheduled Instructional Times That Were Held and Attended	80
Table V-1.	Results for Practical Achievement Scale	86
Table V-2.	Results for Comprehensive Test of Basic Skills (Level 2 Reading Test Grades 4 and 5)	87
Table V-3.	Results for Comprehensive Test of Basic Skills (Level 2 Mathematics Test Grades 4 and 5)	88
Table V-4a.	Proportion Correct on Arithmetic DRT for Students First Tested at Wave 1 in Randomly Selected Institutions . . .	91
Table V-4b.	Proportion Correct on Arithmetic DRT for Students First Tested After Wave 1 in Randomly Selected Institutions . .	92

LIST OF TABLES (Cont'd)

		Page
Table V-5a.	Proportion Correct on Literal Comprehension DRT for Students First Tested at Wave 1 in Randomly Selected Institutions	93
Table V-5b.	Proportion Correct on Literal Comprehension DRT for Students Tested After Wave 1 at Randomly Selected Institutions	94
Table V-6	Mean Proportion Correct and Mean Proportion Omitted on Achievement Tests for Students Whose Performance Decreased From Wave 1 to Wave 4	95
Table V-7.	Arithmetic Performance by Score Group and Program	105
Table V-8.	Language Arts Performance by Score Group and Program	105
Table V-9.	Estimated "True" Arithmetic Performance by Score Group and Program	106
Table V-10.	Estimated "True" Language Arts Performance by Score Group and Program	106
Table V-11.	Performance of Students Who Maintained or Improved Scores Over Time	107
Table V-12.	Responses to Self Esteem Scale	109
Table V-13.	Responses to Locus of Control Scale	110
Table V-14.	Responses to Peer Affiliation Scale	111
Table V-15.	Responses to School Involvement Scale	112
Table V-16.	Responses to Liking of School Scale	113
Table V-17.	Responses to Sociability Scale	114
Table V-18.	Responses to Attitudes Toward Mathematics Scale	115
Table V-19.	Responses to Attitudes Toward Reading Scale	116
Table V-20.	Responses to Teacher Support Scale	117
Table V-21.	Responses to Practical Orientation of Learning Scale	118
Table V-22.	Responses to Expectations for Release Success and Stigmatization Scale	119

LIST OF TABLES (Cont'd)

		Page
Table V-23.	The Relationship of Educational Program Characteristics to the Proportion Correct on the Literal Comprehension DRT at Wave 3 for Students in Programs of Title I Plus Regular Education Instruction	124
Table V-24.	The Relationship of Educational Program Characteristics to the Proportion Correct on the Literal Comprehension DRT at Wave 3 for Students in Programs of Only Regular Education Instruction	124
Table V-25.	The Relationship of Educational Program Characteristics to the Proportion Correct on the Arithmetic DRT at Wave 3 for Students in Programs of Title I Plus Regular Education Instruction	125
Table V-26.	The Relationship of Educational Program Characteristics to the Proportion Correct on the Arithmetic DRT at Wave 3 for Students in Programs of Only Regular Education Instruction	125
Table V-27.	Regression Analysis for Average Self-Esteem Scores of Non-Title I Students at the Third Testing	128
Table V-28.	Regression Analysis for Average Self-Esteem Scores of Title I Students at the Third Testing	128
Table V-29.	Regression Analysis for Average Locus of Control Scores of Non-Title I Students at the Third Testing	129
Table V-30.	Regression Analysis for Average Locus of Control Scores of Title I Students at the Third Testing	129
Table V-31.	Regression Analysis for Average Attitude Toward Reading Scores of Non-Title I Students at the Third Testing . . .	130
Table V-32.	Regression Analysis for Average Attitude Toward Reading Scores of Title I Students at the Third Testing	130
Table V-33.	Regression Analysis for Average Attitude Toward Mathematics Scores of Non-Title I Students at the Third Testing	131
Table V-34.	Regression Analysis for Average Attitude Toward Mathematics Scores of Title I Students at the Third Testing .	131

LIST OF TABLES (Cont'd)

		Page
Table V-35.	Regression Analysis for Average Affiliation Scores of Non-Title I Students at the Third Testing	132
Table V-36.	Regression Analysis for Average Affiliation Scores of Title I Students at the Third Testing	132
Table V-37.	Regression Analysis for Average School Involvement Scores of Non-Title I Students at the Third Testing	133
Table V-38.	Regression Analysis for Average School Involvement Scores of Title I Students at the Third Testing	133
Table V-39.	Regression Analysis for Average Liking of School Scores of Non-Title I Students at the Third Testing	134
Table V-40.	Regression Analysis for Average Liking of School Scores of Title I Students at the Third Testing	134
Table V-41.	Regression Analysis for Sociability Scores of Non-Title I Students at the Third Testing	135
Table V-42.	Regression Analysis for Average Sociability Scores of Title I Students at the Third Testing	135
Table V-43.	Regression Analysis for Average Teacher Supportiveness Scores of Non-Title I Students at the Third Testing . . .	136
Table V-44.	Regression Analysis for Average Teacher Supportiveness Scores of Title I Students at the Third Testing	136
Table V-45.	Regression Analysis for Average Practical Orientation of Learning Scores of Non-Title I Students at the Third Testing	137
Table V-46.	Regression Analysis for Average Practical Orientation of Learning Scores of Title I Students at the Third Testing	137
Table V-47.	Regression Analysis for Average Post-Release Expectation and Stigmatization Scores of Non-Title I Students at the Third Testing	138
Table V-48.	Regression Analysis for Average Post-Release Expectation and Stigmatization Scores of Title I Students at the Third Testing	138

LIST OF TABLES (Cont'd)

		Page
Table A-1.	Description of Language Arts/Reading Domain	A-2
Table A-2.	Description of Mathematics Domains	A-4
Table A-3.	Language Arts/Reading Domain Rankings	A-7
Table A-4.	Mathematics Domain Rankings	A-8
Table A-5.	Possible Item Types for Measurement of Word Recognition Skills	A-10
Table A-6.	Possible Item Types for Measurement of Vocabulary Skills	A-11
Table A-7.	Possible Item Types for Measurement of Literal Comprehension Skills	A-12
Table A-8.	Content Specifications for Language Arts/Reading Domains	A-15
Table A-9.	Item Sampling for Word Recognition and Vocabulary Subtests	A-17
Table A-10.	Distribution by Difficulty Level of Passages from Graded Reading/Literature Texts	A-18
Table A-11.	Number of Multiple-Choice Cloze Exercises by Difficulty Level of Passage, Domain, and Content Area	A-19
Table A-12.	Composition of Literal Comprehension Subtest, by Form	A-21
Table A-13.	Content Specification for Mathematics Domains	A-24
Table A-14.	Content Specification for Addition Subtest	A-27
Table A-15.	Content Specification for Subtraction Subtest	A-29
Table A-16.	Content Specification for Multiplication Subtest	A-31
Table A-17.	Content Specification for Division Subtest	A-33
Table A-18.	Content Specification for Practical Mathematics Subtest	A-35
Table B-1.	Mean Proportion Correct on Arithmetic Subtests for Students Who Achieved Perfect Scores on Later Subtests	B-2
Table B-2.	Intercorrelations Among Addition Subtests at Wave 1	B-3
Table B-3.	Intercorrelations Among Subtraction Subtests at Wave 1	B-3

LIST OF TABLES (Cont'd)

		Page
Table B-4.	Intercorrelations Among Multiplication Subtests at Wave 1	B-4
Table B-5.	Intercorrelations Among Division Subtests at Wave 1 . . .	B-4
Table B-6.	Intercorrelations Among Arithmetic Tests at Wave 1 . . .	B-5
Table B-7.	Intercorrelations Among the Eight Literal Comprehension Passages at Wave 1	B-5
Table C-1.	Factor Scales and Loadings	C-3
Table C-2.	Factor Scale Intercorrelations (below diagonal) and Kuder-Richardson Reliability Estimates (diagonal)	C-5
Table D-1.	Distribution of Student Characteristics Within Institutions	D-2
Table D-2.	Percent of Students in Each Category of Reason for Institutionalization at Each Institution	D-3
Table D-3.	Percent of Students in Each Ethnic Category at Each Institution	D-4
Table E-1.	Summary of the Mean Responses of Teachers of Basic Skills at 30 Randomly Selected Institutions	E-2
Table E-2.	Summary of the Mean Responses of Title I Reading Teachers at 30 Randomly Selected Institutions	E-5
Table E-3.	Summary of the Average Responses of Title I Mathematics Teachers at 30 Randomly Selected Institutions	E-7
Table F-1.	Means and Standard Deviations of Teacher Responses . . .	F-2
Table G-1.	Arithmetic Results	G-2
Table G-2.	Literal Comprehension Results	G-3
Table H-1.	Results of the Analysis of Covariance for Arithmetic . .	H-4
Table H-2.	Results of the Analysis of Covariance for Language Arts .	H-5
Table H-3.	Results of the Analysis of Covariance for Locus of Control	H-7
Table H-4.	Results of the Analysis of Covariance for Liking of School	H-8

LIST OF TABLES (Cont'd)

		Page
Table H-5.	Results of the Analysis of Covariance for Teacher Support	H-9
Table H-6.	Results of the Analysis of Covariance for Practical Orientation	H-10

LIST OF FIGURES

		Page
Figure II-1.	Sequence of Test Administrations	25
Figure V-1.	Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I and Regular Education Students in Randomly Selected Institutions	100
Figure V-2a.	Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I and Regular Education Students at Institution 17	102
Figure V-2b.	Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I and Regular Education Students at Institution 26	103
Figure V-2c.	Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I Students at Institution 7 . .	104
Figure V-2d.	Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I Students at Institution 16 . .	105
Figure A-1.	Passage as Presented in Original Compilation	A-22
Figure A-2.	Passage as Presented in DRT	A-23

EXECUTIVE SUMMARY

INTRODUCTION

This summary presents an overview of the second phase of the National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth. More detailed information is contained in the full report which is Volume 2 of the series entitled Compensatory Education and Confined Youth.

The first phase of the study (reported in Volume 1) provided a description of Title I programs at 100 randomly chosen institutions. The second phase of the study was designed to address the question of whether Title I programs are effective in improving students' basic skills and attitudes, and to determine which components of educational programs and characteristics of institutions are related to such improvements.

Highlighted below are the major findings of this study:

- Scores on standardized tests confirmed that these confined youth, who average 17-1/2 years of age, performed on a par with public school students at about the 5th-grade level.
- The data from the standardized tests indicate that 30 percent of the students who do not receive Title I services are at least as needy as the average student who does receive such services.
- The data suggest that Title I students are offered less total time in reading instruction (65 percent of which is provided by Title I) than non-Title I students. Because Title I students attend less of the instruction they are offered, they receive even less instruction in reading than do non-Title I students.
- The data suggest that Title I students are offered much more time in mathematics instruction (60 percent of which is provided by Title I) than non-Title I students. However, their attendance rate is worse for mathematics than for reading, so they receive only a slightly larger amount of instruction.

- Title I instruction costs more on a per-pupil basis than non-Title I instruction. The costs are higher because more hours of instruction are offered to fewer students.
- Work details, field trips and parole board appointments scheduled by the institutions often conflicted with instructional offerings. When students are punished by being put in "lock-up" they may not receive educational services. Other personal factors such as sickness and home furloughs also reduce attendance in classes.
- In view of the low rate of attendance, it is not surprising that over a twelve-week period of instruction the students' basic skills performance and attitudes did not change. Title I students showed no differential change compared to non-Title I students, on the average.
- Students at some institutions did seem to improve in basic skills or attitudes over time. Institutions that had smaller programs and low rates of staff turnover, that tended to interfere less with educational planning and operation, and that viewed custody as secondary to education or counseling seemed more likely to be successful.

The remainder of this executive summary is organized into two major sections: the first describes the study design while the second provides more detail about the findings highlighted above.

STUDY DESIGN

The study was non-experimental in nature, surveying a nationally representative random sample of 30 institutions and an additional sample of ten institutions judged to be of particular interest. Four visits were made to these institutions, at intervals of six weeks between November of 1977 and April of 1978. During each visit students were tested using specially designed tests focused on basic skills in reading and arithmetic. In addition, the students were given standardized tests at one of the visits in order to compare their performance to that of non-institutionalized populations. Attitude scales were given at the first and third visits. A total of 1,500 to 1,700 students were tested on each visit.

Data were also collected from the administrators of programs, the facility directors and fiscal officers, and the teachers at all 40 sites. These data were used to describe the settings and the educational programs within them and to determine whether the characteristics of institutions or educational programs were related to student improvements in basic skills or attitudes.

Teachers were also asked to keep records of the hours of classroom instruction that the students could have received and actually did receive. Like the site and program descriptions, these records were used to relate the amount of service to achievement gains and attitude changes.

FINDINGS

The principal finding of the study documents the widely-held notion that students in state-run institutions for the neglected or delinquent perform very poorly on tests of basic skills achievement. The performance of institutionalized Title I students, who are 16-1/2 years old on the average, was similar to that of non-institutionalized fourth- and fifth-grade students. Institutionalized students not in Title I (who are 18 years old) performed about one grade level higher. About 30 percent of the students who are not served by Title I programs are achieving at levels below that of the average student who is given Title I services. Apparently in response to the need for services, nearly one-half of the institutions report that they provide some Title I services to all of their eligible students.

The second major finding was that the achievement levels and the attitude measures remained stable across the 12 weeks between the first and third testing. On the average, neither the non-Title I students nor those in Title I supplementary programs showed any meaningful change in these measures over this time period. Although this result is disappointing, it should be recalled that the non-experimental nature of the study makes it difficult to know how Title I students would fare if they did not receive supplementary services. Perhaps the scores of these students would decline if they could not receive educational services at their level.

With regard to the services provided by these programs the data on instructional time, while incomplete, indicate that Title I students are offered less instructional time in reading, but more instructional time in mathematics than are non-Title I students. Title I students are offered about three hours per week of Title I reading instruction, and one and one-half hours per week of regular educational reading instruction. The non-Title I students are offered four hours per week of regular reading instruction, and they also are offered nearly one hour per week of Title I instruction. Because of poor attendance (described in more detail below) the Title I students actually receive only about three hours per week total instruction in reading compared to nearly five hours per week actually received by the non-Title I students.

In mathematics, the students in Title I are offered a total of seven hours per week of instruction (four hours of this is provided by Title I). Non-Title I students are offered almost three hours per week, none of which is provided by Title I. Here, the attendance rate for Title I students (especially in the Title I classes) is very low, resulting in Title I students receiving only two and one-half hours per week of mathematics instruction compared to a total of two hours per week for non-Title I students.

The large gap which seems to exist between the services offered and the services received seems to be due to two causes: institutional priorities interfering with the educational program and student unwillingness to attend classes. The institutional interference can affect groups of students or individuals. The effect on groups is seen most clearly in the loss each week of nearly one-half hour out of nearly three hours of scheduled Title I instruction in reading because the classes weren't held. Although the reasons for this loss are not known, one can speculate that the teacher was absent and could not be replaced or that all the students in the class were on work release, a field trip, or in "lock-up".

These events can also happen to individual students, preventing them from attending the classes which are held. Analyses of the data obtained in this study provide a tentative indication that about 14 percent of the students

on a typical day are either sick, locked up (for disciplinary reasons), have conflicting appointments (e.g., with a parole board), are on work release, are on furlough, or have gone AWOL. An unknown number of other students simply don't attend the classes which are given.

The costs of the Title I supplementary services are higher, by 50 to 100 percent per pupil, than the costs of the regular (state-supported) education programs in reading and mathematics. The Title I program provides more instructional time, as indicated above, and it serves fewer students. The gap between the services offered and those received seems especially important when the costs of the services are considered.

The remaining findings concern the relationships of characteristics of institutions and educational programs to achievement gains and changes in attitudes. Because many characteristics were included in these analyses, what follows represents a considerable degree of generalization from the data.

One important generalization from these analyses is that the relationships of institutional and educational program characteristics to the achievement gains and attitude changes are consistent across programs. That is, a characteristic which is positively related to achievement gains or attitude changes for Title I students is also positively related to such outcomes for non-Title I students. This is important because it means that changing the characteristics of educational programs or institutions should have similar consequences for both programs.

In general, it seems reasonable to conclude that institutions which are able to devote time and resources directly to programs aimed at the achievement and attitudinal needs of their students are more likely to be successful than others in effecting changes in their students. Specifically, smaller programs, with low rates of staff turnover, in institutions that consider custody to be less important than education or counseling, and where the institution does not interfere in the planning or operation of the educational program, are more likely to be successful.

CHAPTER I. INTRODUCTION

BACKGROUND AND OBJECTIVES OF THE OVERALL STUDY

To expand and improve the education of confined youth, Congress has provided funding for basic skills programs in institutions for neglected children and delinquent youth, and in adult correctional facilities housing youthful offenders. This funding stems from Title I of the Elementary and Secondary Education Act of 1965, amended in 1972 to include youths institutionalized in adult correctional facilities. Title I guidelines state that the funds should be used to provide services to residents who are under 21 and who have not received high school diplomas. These services should concentrate on basic reading and math skills, and should supplement, rather than supplant, other educational programs in the facility.

The objectives of Congress in providing these funds are, in general, to help the institutionalized population build basic skills and to overcome the effects of prior academic failure. These Title I funds have been available to state institutions housing neglected or delinquent youth since 1967, as authorized by P.L. 89-750, passed November 3, 1966.

The National Evaluation of Title I Programs for Neglected or Delinquent Youth in State Institutions was mandated by Part (a) of Section 151 of P.L. 89-10. Under contract to the Office of Evaluation and Dissemination of the U.S. Office of Education (USOE), System Development Corporation (SDC) began an evaluation in February 1976 with three initial major objectives. The first objective, accomplished during Phase I of the project, was to present descriptions of representative Title I programs in state institutions. A report containing the results of the Phase I study has been submitted to USOE and has been distributed to state personnel, institutional staff, members of Congress and other interested parties.

The second objective of the evaluation was to develop evaluation models and reporting forms that can be used by institutions and state agencies for annual evaluation reports to the U.S. Office of Education. SDC has reviewed reports

of Title I N or D program evaluations available to USOE. On the basis of this review--and the experience gained during the course of the national evaluation--reporting models have been recommended that will provide standardized information to USOE on program operations and the degree of success realized in meeting Title I objectives. The ultimate goal of this model development activity is to simplify and standardize evaluation and reporting procedures so that federal, state, and local program administrators can obtain information on Title I programs and their nationwide impact. A report describing the recommended models--which also presents data collection forms and instructions--has been submitted to USOE.

The third objective was to measure the impact of the Title I program on student acquisitions of basic skills, i.e., language arts, and mathematics, and on the attitudes of program participants about themselves, about school, and about learning in general. This volume presents the findings of Phase II of the study: the impact evaluation.

PHASE I OBJECTIVES

The U.S. Office of Education wanted to obtain information about the ways in which Title I programs have been implemented nationwide. Phase I of the study was designed to describe Title I programs with respect to:

- The proportion of residents who are eligible for Title I services and the proportion of eligibles served
- The estimated skill levels of the residents served
- The length of time residents stay in the program
- The subject matter content of the Title I program
- The institutional settings in which Title I programs operate
- The management of the programs
- The amounts and allocations of funding to support the Title I program

OVERVIEW OF PHASE I FINDINGS

In order to gather information for the program descriptions in Phase I, SDC sent two-member teams of trained interviewers/test administrators into the field. Nationwide, these teams visited 100 randomly selected facilities for delinquents, adults, and neglected children* that receive Title I funds. They also interviewed state educational and correctional administrators of Title I programs in each state. An overview of major findings is presented in the following paragraphs.

Characteristics of the Students Served by Title I Programs

About one-half of the sample institutions reported that they serve 50 percent or fewer residents eligible for Title I services. Of the three types of facilities, adult facilities have the lowest proportion of participation in the program, the greatest number of educationally needy residents not being served, and the most difficulty in maintaining a constant level of Title I program services. The number of eligible residents in adult facilities has increased by 125 percent over the past three years, while the number of eligible residents in delinquent and neglected facilities has remained relatively stable.

With respect to student performance, teachers estimated that 50 percent of the Title I students in neglected facilities were reading at grade levels 1-3, as were 40 percent of the students in delinquent facilities and 25 percent of the students at adult institutions. Less than one third of all Title I reading students were perceived as functioning at a 6th-grade level or above, with adult institutions reporting the greatest proportion of students in this group (38 percent, as compared to 26 percent in delinquent facilities and 8 percent in neglected facilities).

Over 60 percent of the Title I students in N or D facilities are in the Title I program for six months or more. The students who receive the longest exposure to Title I instruction are those in facilities for the neglected.

*Facilities for neglected children were over-sampled in Phase I only in order to allow for a description of them as a type.

Content of the Title I Program

Although reading and math are by far the most heavily emphasized content areas in Title I programs, a variety of other areas is being taught as well. Within the two major instructional areas, reading instruction emphasizes vocabulary, literal comprehension, and following directions, while the emphasis in math is on fundamental operations, practical mathematics, and word problems. Classroom activities are largely structured, with students spending much of their time working independently on assigned activities or working individually with the teacher. Materials most commonly used in reading instruction are written materials and student worksheets, although other types of materials are also used frequently. In math, however, worksheets and audiovisual equipment are by far the most commonly used materials.

Title I programs spend a higher percentage of their funds on inservice education and place more emphasis on individualization of instruction and assessment of student progress than do regular education programs. On the other hand, regular education programs place less emphasis on reading and math, and more emphasis on vocational education, than do Title I programs.

Title I teachers were generally supportive of students, and interaction between teachers and students was frequent. Teachers typically responded to student errors with constructive criticism, and students generally were reported to be open in their responses to their teachers.

Title I Setting

Sites in the study sample were examined to learn something about the contexts in which Title I programs operate. Institutions ranged in size from 11 to 2500 residents. Institutions for the neglected reported larger resident populations, on the average, than institutions for delinquents; 70 percent of the neglected facilities, and 47 percent of the delinquent facilities, reported populations of more than 100 residents.

In neglected and delinquent facilities, residents were usually housed in small cottages, in which custody and security precautions were minimal. In adult facilities, residents were usually housed in large dormitories or individual cellblocks. Security restrictions were generally the same in education programs as they were in the rest of the institution; only in adult institutions (where education programs tend to be highly selective of participants) were educational program staff allowed to exercise fewer security restrictions than did the facility as a whole.

State institutions with Title I programs appear to vary dramatically on most of the institutional and population descriptors. This variation was found to be related to the custody or treatment orientation of the institution. The over-all resident/staff ratio for the average institution was 3:1, but most of these staff members occupied custody or central administration roles. The average institution reported a ratio of one treatment and one education staff member for every thirty-two residents, with the number of treatment and education staff being considerably larger in treatment-oriented facilities.

Custody-oriented institutions were characterized by their larger size, their overcrowdedness and their higher proportions of custody staff. Custody orientation was associated with fewer education and treatment staff members per resident. Regular education and Title I programs in custody-oriented institutions placed more emphasis on vocational education than did education programs in treatment-oriented institutions. Education programs in treatment-oriented institutions emphasized affective growth more than did programs in custody-oriented institutions.

Program Management

The ways in which Title I programs were managed in different institutions were examined and compared. The topics covered included effects of federal guidelines, monitoring and evaluation, and patterns of decision-making.

There is at least some knowledge of federal guidelines among almost all Title I state and facility administrators. The most widely known guidelines appear to

be (1) that Title I should supplement existing education programs, rather than supplant them; (2) that students must be under 21 years old to receive Title I services; (3) that Title I should serve the educationally deprived, and (4) that Title I should provide basic skills instruction in reading and math.

The primary means by which information about guidelines is disseminated is through the state education agencies (SEAs). Title I administrators express considerable dissatisfaction about the lack of clarity of federal guidelines. Many administrators believe that guidelines are not being closely followed, either in their own state or across the country, even though the USOE-recommended program components exist in some form at almost every institution. When state coordinators were asked what changes they would recommend in federal guidelines, the most common responses dealt with the elimination of specific requirements concerning student eligibility or program content. Frequently suggested changes were to provide subjects other than reading and math, to eliminate age restrictions, to expand eligibility beyond the most needy so that all needy could be served, and to remove the restriction that Title I must supplement existing programs.

The vast majority of Title I projects are monitored and evaluated by state agencies. Monitoring is used to assess compliance with guidelines and to assess the need for changes in the program and in state assistance. However, state and facility administrators expressed considerable dissatisfaction with the quality of current project monitoring and evaluation.

State applicant agencies (SAAs) tend to be involved in decision-making about more aspects of program operations than are state education agencies. SAAs (who are generally departments of corrections for adults or youths) are more involved than SEAs in decisions on matters related to the facility or classroom level, such as defining the neediest students, setting goals for instruction, assigning resources for classes, and selecting basic instructional materials. On the other hand, the SEA appears to be more involved than the SAA in decisions on the level of funding and interpretation of federal guidelines.

Title I Funding

Information on funding was gathered to describe the amount and allocations of Title I funds used at the facility level to provide services to students. Generally, the largest single portion of the state or facility education budget came from state sources. Title I funds are the next single largest contributor. Little Title I money was retained by the state agencies; in fact, on the average, over 93 percent of the Title I money was distributed to the participating facilities.

The average Title I student in the sample facilities received $1358 in services from the combination of all funding sources. Of this amount, approximately $456 was derived from Title I funds. These figures varied considerably across institutions. Statistically significant differences in the level of state per-student expenditures, and total expenditures per Title I student, were found between different types of institutions. Adult institutions received, on the average, half the amount of state funding per student that neglected facilities received. The average total per-student expenditure was $2002 in institutions for the neglected, $1469 in delinquent institutions, and $905 in adult institutions.

The major portion of Title I and state education money was allocated to the basic skills area. State education funds, and other federal funds, were used extensively for vocational education. Of the funds spent for educational services, the largest portion (over 70 percent of Title I and 80 percent of state funds) went for personnel. Most of this money was spent for classroom teachers and aides, particularly in the case of Title I funds. The state money was the primary support for administrative services. The money not spent for personnel was channeled primarily into instructional materials.

Opinions Concerning the Operation and Effectiveness of Title I Programs

In the course of the interviews, opinions were solicited on the strengths, weaknesses and effects of the Title I program. Among teachers and administrators, the strongest aspects of the Title I program were perceived to be individualized instruction and effectiveness in improving basic skills. Other aspects

mentioned frequently were the program's qualified staff, good equipment and materials, and the small size of instructional groupings. The most frequently mentioned weaknesses of the Title I program were the restrictiveness of guidelines and insufficient funds.

The majority of state coordinators reported some problems in enforcing guidelines, particularly the "supplementing rather than supplanting" requirement. Facility personnel, on the other hand, reported problems that resulted from the successful enforcement of this guideline. Specifically, the interpretation of this guideline would result in a duplication of services between the Title I and regular education programs; or overconcern that participants, resources, and services be strictly distinct.

Staff members in all roles reported that there was poor communication between teachers and facility staff and administrators. Lack of communication among staff members was a source of much concern--a concern consistent with the fact that residential institutions (where resident behavior must be constantly monitored) afford little opportunity for staff interaction. Reported conflict between education and custody staff took the form of custody staff resentment of the intrusions of educational programs into correctional institutions. Such conflicts may reflect the differing priorities assigned to custodial and educational goals.

Students reported a generally high level of interest and involvement in the Title I program, with students in facilities for the neglected being somewhat less positive in their ratings than students in delinquent and adult institutions. Students in adult facilities were the most positive about how much they were learning in Title I classes and about the value they placed on these educational activities. Good teaching methods, personal attention, and adequate materials were reasons given for positive feelings about the Title I program by respondents in all types of facilities. Students appeared to be satisfied with the way they were being taught and with what they were learning. Those who suggested changes in the Title I program asked for a greater variety of classes and more intense exposure to educational activities.

Based on perceived effectiveness ratings, most Title I administrators and teachers think that the Title I program is improving the educational performance of participants. Most respondents reported positive spin-offs of the Title I program to other areas. The expanded use of needs assessment throughout a facility's educational programs, increased teacher awareness of new teaching methods, more positive staff attitudes toward remedial education, and increased student self-esteem were all cited as by-products of Title I program operations. Whether these perceptions are reflective of program effectiveness is the subject of this Phase II report.

PHASE II OBJECTIVES

Phase II of the study was designed to answer these specific policy questions:

- Is Title I effective in improving student proficiency in basic skills?
- What are the effective components of Title I programs?
- Are different components of Title I programs more effective in some kinds of sites than in others?
- Are different components of Title I programs more effective for some kinds of students than others?
- Is improvement in basic skills related to a student's school-related attitudes?

BRIEF GUIDE FOR THE READER OF VOLUME II

The remainder of this volume is organized into four chapters and several appendices. Most of the technical details and the lengthy descriptions of data are to be found in the appendices. Each chapter begins with a summary of the major points in the chapter.

Chapter II describes the design of the Phase II study. Chapter III describes the instruments used in the study. Chapter IV describes the general characteristics of the sample of students, the educational programs, and the institutional environments. Chapter V reports the results of the impact study and the various analyses performed to relate program characteristics to outcomes.

CHAPTER II. DESIGN OF PHASE II IMPACT EVALUATION

PRECIS OF CHAPTER II

This chapter presents the design for the Phase II study. The study was designed to obtain information about the effectiveness of Title I programs. Specific information about the components and operations of Title I and state-supported educational programs was to be related to the changes in attitudes and in basic skills achievement observed over the 18 weeks of the study. The study was designed to gather information about all facets of the institutions and their educational programs from a nationally representative sample of 30 institutions--and from an additional 10 institutions judged by SDC to be of particular interest.

In order to assess changing performance levels on tests of basic skills, the institutions were visited four times at six-week intervals. During each visit, students were tested using SDC-developed domain referenced tests of basic skills. The six-week interval was chosen to maximize the chances for obtaining repeated testings of students while allowing time for instructional effects to occur.

Between 1500 and 1700 students were tested during each visit. An important finding concerning the attrition of students from the study was that, of those students tested on the first visit, nearly one third were still enrolled in educational programs at the fourth visit, but were not available for testing. An analysis of the reasons given for not being present for testing at the second visit indicated that 72 percent of these students either were reported to be sick, were placed in "lock-up", or were busy with other activities which created scheduling conflicts. This indicates that there are many interruptions to student participation in educational programs in these institutions.

This chapter gives additional detail on the sampling plan, the use of data collection instruments during each visit, and the field operations undertaken by SDC personnel.

CHAPTER II

OVERVIEW OF THE STUDY DESIGN

The major objective of the work reported in this volume was to measure the effects of the Title I program on pupil achievement in the basic skills of reading and mathematics, and on pupil attitudes. The specific policy questions to be answered were:

- Are Title I programs effective in improving student basic skills?
- What are the effective components of Title I programs?
- Are different components of Title I programs more effective in some kinds of sites than in others?
- Are different components of Title I programs more effective for some kinds of students than others?
- Is improvement in basic skills related to a student's attitudes toward school?

The study was non-experimental in nature, surveying current practices in state institutions across the nation. In order to address the questions stated above, the study was designed to gather information about all facets of the institutions and their educational programs for a nationally representative sample of 30 institutions and from an additional 10 institutions judged by SDC on the basis of Phase I observations to be of particular interest.

The greatest effort was expended upon the assessment of student performance in basic skills. Institutions were visited four times at six-week intervals and students were tested each time. The six-week cycle for visits was formulated on the basis of two factors that are important to the evaluation of Title I programs in neglected or delinquent institutions: (1) test administrations should be scheduled at intervals of sufficient length that gains in achievement can be detected; and (2) testing intervals should be relatively short to enhance the availability of sample students for subsequent test administrations.

Between 1500 and 1700 students were tested during each visit using SDC-developed tests of performance in basic skills. Other instruments were used at each visit

also. Table II-1 provides a guide to the use of instruments in this study. Institutional staff maintained records of dates of student entry to and exit from educational programs, as well as the time students actually spent in instructional settings.

Teachers, educational program administrators and facility directors provided information about the goals and operational characteristics of the educational program and the Title I components in particular. Information on the costs of the educational program was obtained from each institution's fiscal officer. Table II-1 indicates the content of the instruments used during the Phase II study. Chapter III gives more information on the content of these instruments.

Two sets of analyses were planned:

1) The performance tests in basic skills were to be used in a direct assessment of the impact of the Title I program on student performance in basic skills.

2) The relationships between the characteristics of the programs and student performance in basic skills were to be examined.

Chapter V presents the results of these analyses.

SAMPLING PLAN

The sampling plan for this study was based upon knowledge of project characteristics discovered during Phase I.* The plan was reviewed and revised according to suggestions made by the Policy and Research Advisory Groups.

Two variables were used for stratification: degree of emphasis on basic skills, and average per-student expenditure. The Phase I study indicated that Title I programs in institutions for neglected or delinquent youth could be characterized

*Bartell, T. et al., Compensatory Education and Confined Youth: A National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth. Volume 1. Santa Monica: System Development Corporation, TM-5792/005/00, 1977.

Table II-1. Content of Instruments Used in the Phase II Study

Instrument	Content
Facility Director Questionnaire (FDQ)	Characteristics and goals of the institution and the educational program
Facility Fiscal Questionnaire (FFQ)	Costs of operating the regular education and Title I instructional programs.
Educational Program Administrator Questionnaires (EPA)	Characteristics and goals of the educational programs
Teacher Questionnaires	
Part A (TQA)	Characteristics and goals of the educational programs
Part B (TQB)	Implementation and process information about the Title I reading program
Part C (TQC)	Implementation and process information about the Title I mathematics program
Student Questionnaire (SQ)	Student background information
Student Data Sheet (SDS)	Record of student entry into and exit from educational programs
Student Participation and Attendance Record (SPAR)	Record of student time in instructional settings
Classroom Observations (CO)	Data collectors' assessments of program characteristics and goals
Test Session Observations (TO)	Data collectors' record of testing activities
Attitudes Toward School and Learning (SAA)	Paper and pencil survey of student attitudes
Practical Achievement Scale (PAS)	Paper and pencil test of student achievement related to practical settings
Domain Referenced Tests (DRT)	Paper and pencil tests of student achievement in basic skills
Comprehensive Test of Basic Skills (CTBS)	Standardized, norm-referenced, paper and pencil test of achievement

as providing instruction of one of the following types: 1) instruction in reading and math only; 2) instruction emphasizing reading and math; 3) instruction that placed a minor emphasis on reading and math. These program types constituted the index of "degree of emphasis on basic skills". Title I per-student expenditures were used as the other stratification factor to assure that the full range of such expenditures was included in the study.

Some of the Phase I facilities were excluded from this new sampling frame due to special problems that would make student testing impractical. Sample facilities in Puerto Rico (N=3), where only Spanish is spoken, were dropped from the sampling frame in order to eliminate problems of ensuring comparability of test items across languages. Three sites with programs that lasted less than four weeks were eliminated because such short-term programs are likely to evidence effects too small to be detected, and few students would be available for follow-up testing. Three more sites were dropped from the frame because they were no longer operating Title I programs.

Since the data from institutions for the neglected indicated only slight program differences between programs for the neglected and programs for the delinquent, the former were undersampled by a factor of 0.56 for this study to equalize their overall sampling fraction with that of adult and delinquent institutions*--thus producing a self-weighting sample for our analyses. This procedure resulted in approximately 6 of the 10 institutions for the neglected in the Phase I study being eligible for selection in this study.

The remaining 87 sites were distributed as indicated in Table II-2. Thirty of these sites were selected with equal probability so as to allow unweighted comparisons across program types. This approach yielded the numbers of sites by cell, as indicated in Table II-3.

---

*Neglected facilities were oversampled by a factor of 1.79 in the earlier study. Undersampling selected facilities by a factor of 0.56 produces an overall sampling fraction for these sites in the impact study equivalent to that of delinquent and adult facilities (1.79 X 0.56 = 1.00).

Table II-2. Number of Facilities Eligible for Selection into Phase II Samples, by Strata

Title I Program Emphasis	Title I Per-Student Expenditure $501+	$201-500	$1-200
Reading, Math Only	10	7	10
Strong Emphasis on Reading, Math	11	13	8
Weak Emphasis on Reading, Math	12	8	8

Table II-3. Number of Facilities in Phase II Sample, by Strata

Title I Program Emphasis	Title I Per-Student Expenditure $501+	$201-500	$1-200
Reading, Math Only	3	2	3
Strong Emphasis on Reading, Math	4	5	3
Weak Emphasis on Reading, Math	4	3	3

In addition to the 30 sites sampled with strict probability methods, 10 additional sites were purposively selected for their likelihood of being effective. Criteria measured in the Phase I study, such as strong staff commitment to the program, staff cohesiveness, high quality administrative leadership in the educational context, positive student and staff rapport, and the utilization of innovative approaches were used to select these sites.

Individual students within sample facilities were selected with probability methods. The goal was to test 35 Title I students in each site where that number of students was available, and to test all Title I students in smaller programs. Comparison students from other remedial and regular academic education programs were sampled and tested in those sites where such programs existed.* Students participating only in post-secondary education programs or only in vocational education programs were eliminated from the sampling frame. The intent was to test a total of approximately 25 comparison students in sample institutions where that number of students was available. In the smaller facilities and programs, all residents were tested at each point in time to minimize disruptions to these facilities caused by the selection of all but a very few students.

One important implication of not taking a constant proportion of students from each sample facility is that, while institutions were selected with equal probability, the total set of individual students did not constitute a self-weighting sample. When students were pooled across facilities to do individual-level analysis, weighting factors had to be introduced to correct for unequal selection probabilities.

---

*It should be pointed out that the designation of students as Title I students or as recipients of Title I services was based on differing criteria among the institutions. In some institutions Title I students were those assigned to individualized labs; in other institutions they were identified on the basis of their receipt of instruction from teachers whose salaries were paid by Title I funds; in still other institutions Title I students were those students utilizing materials and equipment paid for with Title I funds.

The weighting factors range from 1.0 to 5.2 across the 30 random sites for Title I students, and from 1.0 to 8.3 across sites for non-Title I students. At 20 of these sites all Title I students were tested, while at 21 of these sites all regular education students not receiving Title I instruction were tested. In addition, one of the randomly sampled sites provided Title I services to all students. Thus, only eight of the 30 randomly sampled sites had weights other than one (1.0) for the non-Title I students.

One of the major problems in collecting data from these institutions was the variable availability of students for testing. It was anticipated that the constant release of students from the facilities would reduce the number of students available for the planned four waves of testing. In order to compensate for this attrition, lists of alternate students from the first rosters were provided during the first wave of data collection. In institutions where sampling was performed, substitutions were made from those lists in order to retain, as nearly as possible, 35 Title I and 25 non-Title I students. These students were then maintained in our sample throughout the study. No more selections were made from students who were entered on the first rosters.

To maintain a representative sample across testings, students who entered an educational program between testings were added to the roster and selected at the same rate as previously tested students (per program, per institution) and retained for subsequent test administrations. A subsequent section of this chapter examines the retention rate of students tested in the first visit.

FIELD OPERATIONS

The field operations staff performed three major tasks during Phase II of this study. First, they were responsible for communication between SDC and the 40 institutions participating in the study; second, they selected and trained the data collectors who conducted the actual institution visits; and third, they performed all the tasks associated with coordinating and scheduling the arrival of personnel and materials at each institution. In addition to these major tasks, the field operations staff monitored the activities of the data collectors during each of the four data collection trips.

During the fall of 1977, Dr. John W. Evans, Assistant Commissioner of USOE's Office of Planning, Budgeting, and Evaluation (OPBE) wrote to all State Administrators of the ESEA Title I Programs in institutions for neglected or delinquent children. His letter explained the second phase of the study and enclosed the list of selected institutions. A second letter was then sent to the 40 institutions from the project director explaining in more detail the intent of the Phase II study and the basic involvement of each institution.

The field operations staff then called the Phase I contact person at each institution to provide more information about the Phase II study, specifically regarding the test and data collection activities scheduled for that institution. During this preliminary discussion the staff members accomplished the following objectives:

- The institutional contact person for the Phase II evaluation was identified.
- A broad overview of the institution's task involvement in Phase II was presented.
- Dates and schedules for the first visits were established.

After each institutional visit was scheduled, a Procedures Manual for institutions participating in the second phase of the impact study was mailed to each contact person. Table II-4 shows the content of this manual.

Table II-4. Contents of the Procedures Manual

• Background and objectives of the study
• Purpose of the Phase II site visits
• Procedures for assuring the confidentiality of student data
• Student selection procedures
• Site preparations prior to visits
• Site activities during visits
• Site activities between visits

Enclosed with the Procedures Manual was a set of blank student rosters used to identify the population of students from which the participants in the study would be selected. Prior to each subsequent visit the institutions were provided new rosters upon which to list students who had entered the education program since the completion of the previous roster.

These rosters were designed so that the contact person could link student names to identification numbers. Only the student identification number and program data were sent to SDC. In this way student anonymity was ensured. These rosters were then used to aggregate students by Title I and regular education program types and to select those to be tested.

Data were collected at each participating facility four times during Phase II of the study. Each visit required two days and was conducted by a two-person team. Eight data collectors, working in two-person teams, were able to visit the 40 sites within a four-week schedule. Institutional visits were scheduled at six-week intervals, as follows:

Site Visit	From	To
1	November 28, 1977	December 23, 1977
2	January 9, 1978	February 3, 1978
3	February 20, 1978	March 17, 1978
4	April 3, 1978	April 28, 1978

During the two-day visit, each data collection team was responsible for three activities:

- Questionnaires

  The team was involved in administering, distributing, and assisting institutional staff in the completion of six questionnaires and two institutional reporting forms.

- Observations

  During each visit the team recorded their observations about test administrations. During the second visit they recorded observations about the institution and the educational program.

- Tests

  The team administered achievement tests, attitudinal measures, and questionnaires to students during the four visits to the sites.

Table II-5 shows the schedule of activities during each site visit.

Prior to the arrival of the data collection team at the institution, the field operations staff mailed all testing materials to the institution. Also prior to the team's arrival, the institution scheduled students for testing during the two-day visit. The following guidelines for test scheduling were used: test sessions should not have more than fifteen students each, and each student should be scheduled for two different sessions during the visit. If possible, the two test sessions for a student were to be scheduled on different days. The first activity for data collectors upon arrival at an institution was to confirm that all of the pre-visit activities had in fact been accomplished. During the first test session the data collector explained to students the purpose of the testing, how they were selected, confidentiality procedures, and the voluntary nature of the testing. The general procedure followed by the data collectors was to read all items and responses--unless reading was part of the tested outcome. Only the Comprehensive Tests of Basic Skills required a timed administration. The reading aloud of items on other tests did tend to set a pace, however. The first session usually required about one and a half hours from the time the students entered until they left; the second session was usually shorter.

After the test administration, the team collected, packaged and mailed all materials to SDC. Figure II-1 displays the order in which tests were administered during the four site visits.

RATES OF STUDENT PARTICIPATION IN TESTING

As noted earlier in this chapter, the site visitation schedule was designed to maximize the number of students available for subsequent testings as well as the amount of change which could occur between testings. It was anticipated that there would be attrition of students across testings due to releases from

Table II-5. Schedule of Data Collection Activities for Phase II Study

ACTIVITY	VISIT			
	1	2	3	4
Questionnaire Administration or Distribution				
Facility Director (FDQ)	X*			
Facility Fiscal (FFQ)	X			
Educational Program Administrator (EPA)	X			
Teacher Questionnaire				
Part A (TQA)	X			
Part B (TQB)	X			
Part C (TQC)	X			
Student Questionnaire (SQ)	X			
Student Data Sheet (SDS)	X	X	X	X
Student Participation and Attendance Record (SPAR)	X	X	X	X
Field Observations				
Classroom Observations (CO)		X		
Test Session Observations (TO)	X	X	X	X
Test Administrations				
Attitudes Toward School and Learning (SSA)	X		X	
Practical Achievement Scale (PAS)		X		
Domain Referenced Tests (DRT)	X	X	X	X
Comprehensive Tests of Basic Skills (CTBS)				X

*An "X" indicates that this instrument was administered during this visit.

SESSION	INSTITUTION VISIT			
	FIRST	SECOND	THIRD	FOURTH
I	Student Questionnaire (All Students) Attitudes Toward School and Learning Domain Referenced Test Mathematics	Student Questionnaire (New Students) Practical Achievement Scale Domain Referenced Test Mathematics	Student Questionnaire (New Students) Attitudes Toward School and Learning Domain Referenced Test Mathematics	Student Questionnaire (New Students) Domain Referenced Test Mathematics Domain Referenced Test Reading
II	Domain Referenced Test Reading	Domain Referenced Test Reading	Domain Referenced Test Reading	CTBS

Figure II-1. Sequence of Test Administrations

the institutions and the completion of educational programs. Table II-6 presents the percentage of students selected for the first testing who were still at the institution for subsequent testings.

Table II-6. Percentage of Students Selected for First Testing Who Were Present at Institutions for Later Testings

Visit	Percent Students Remaining
Second	83
Third	74
Fourth	67

These percentages represent the students who had not been released from the institutions. However, physical presence at the institution was not the only factor determining a student's availability for subsequent testing.

Table II-7 shows the breakdown of initially-selected students tested during the four visits.

Table II-7. Students Selected for Testing During Wave 1 Who Were Tested at Subsequent Site Visits

	FIRST	SECOND	THIRD	FOURTH
Selected for testing*	1265	__**	__**	__**
Tested	1222	825	604	413
Not tested				
Not at institution	0	218	324	414
At institution	43	222	337	438

*Data is for the 30 randomly selected institutions.

**Only those students initially selected on the first roster for participation are considered here.

By the fourth visit, 33 percent of the students originally selected to participate had left the institution. Thirty-five percent of the original group were still at the institution at the fourth testing, but for a variety of reasons they did not attend the test sessions, as noted above. In fact, over half of the students expected at test sessions--but who did not appear--were still at the institution.

An analysis of the Test Session Observations for the second visit indicated that, of the students who were at the institutions but were not available for testing:

- 28 percent were in "lock-up";
- 21 percent were reported to be sick;
- 23 percent had a conflicting appointment which could not be broken (parole board, work release, furlough, etc.);
- 28 percent refused to take the test.

Of the students recorded as no longer at the institution:

- 74 percent had been transferred from the institution;
- 13 percent had left the educational program;
- 13 percent were AWOL from the institution.

These data are presented to show the types of problems faced in collecting complete data on these students. These data should not be treated as true estimates of the magnitude of these problems, because reasons for absences were recorded somewhat sketchily. Indeed, one large institution did not explain the unavailability of some 39 previously tested pupils. The data from subsequent testings is not as reliable, in our estimation, because students and institutions could have learned how to provide an excuse other than refusal when, in fact, a student refused to take the test.

These data provide a benchmark for assessing the amount of interference that educational programs may experience in institutional contexts due to competing institutional activities that may be given priority over educational activities.

CHAPTER III. INSTRUMENTATION FOR THE PHASE II STUDY

PRECIS OF CHAPTER III

This chapter describes the content of the instruments used to collect data during Phase II. In the first stage of this work the instruments used in SDC's study of the Sustaining Effects of Compensatory Education and SDC's Study of the Emergency School Aid Act were reviewed and modified in the light of information obtained in Phase I about educational programs in correctional institutions.

Several instruments were used to collect information about students, institutions, and educational programs. Information was gathered on the amounts of Title I and state-supported educational funding expended on staff, training, equipment and materials for programs in reading and mathematics. Descriptions of participating students and the selection procedures used to assign students to educational programs were obtained. In addition, descriptions of the size of the programs, the quality of the staff and resources, and the operation of the programs were gathered. Teachers were asked to provide details on the methods of instruction, the degree of individualization of instruction, resources utilized, and the availability of support services and personnel. They were also asked to record participation data for each of the sampled students during a typical week prior to each of the visits. Finally, the data collection teams filled out an observation form which recorded their impressions of the operation of the various programs.

Because assessments of student attitudes and achievement levels were important objectives of this study, much attention was given to the selection and development of appropriate measures in these areas. Chapter III explains the process used in developing the domain referenced tests (DRTs) which were used to assess growth in math and language arts achievement over the four testing periods.

The arithmetic tests developed covered the basic operations of addition, subtraction, multiplication and division as applied to whole numbers, decimal

numbers and fractions. Additionally, a practical mathematics test was developed to assess achievement related to practical applications of mathematics. In language arts, tests were developed to measure achievement in word recognition, vocabulary, and literal comprehension. All of the DRTs had desirable psychometric properties, indicating that they are the results of a valid strategy for instrument development.

The development of the attitudinal measures involved an extensive review of the literature and the culling, rewording, and generation of items for inclusion in the questionnaire. A factor analysis of the data obtained from this instrument was performed which resulted in the identification of several dimensions of interest. Self-attitude dimensions obtained were self-esteem, locus of control, sociability, and expectations for post-release success or stigmatization. School-related-attitude dimensions obtained were school involvement, and liking of school; yielded also were two separate dimensions of attitudes toward reading and mathematics. The psychosocial perceptions of institutional environments fell along dimensions labeled peer affiliation, practical orientation of learning, and teacher support. Appendix C presents the detailed results of this factor analysis.

CHAPTER III.

OVERVIEW OF INSTRUMENT DEVELOPMENT

This section describes the data collection instruments used in Phase II. Items included in these instruments were formulated in the following manner. Analytic questions to be answered were specified, and a list of variables addressing each analytic question was then prepared. Previous evaluation instruments used in SDC's ongoing Study of the Sustaining Effects of Compensatory Education on Basic Skills (SES) and in its past study, the Emergency School Aid Act (ESAA), were examined in order to locate items that measured variables of interest. These items were rewritten to render them appropriate to the correctional education context. Where no existing items could be located appropriate to a variable, new items were prepared. Initial drafts of instruments were reviewed by the study's analysis staff, the Policy Advisory Group, and the Research Advisory Group.

Instruments were designed to provide the following general categories of information:

- descriptive information about the education programs and the institution;
- institutional expenditures for educational activities, particularly those related to reading and mathematics;
- background information on students, educational staff, and the institutional staff that frequently interacts with students;
- instructional characteristics of reading and mathematics programs;
- measures of student, teacher and administrator attitudes on a variety of constructs likely to affect program success;
- information on the type and duration of instruction, by student, for each class attended;
- measures of student school-related attitudes;
- measures of student perceptions of the institutional environment;
- measures of student attitudes about themselves;
- measures of student achievement.

After initial drafts of questionnaires and data collection forms were developed, a field test was conducted to identify the problems respondents might have in completing instruments and in providing the data needed to implement the research design. Respondents were asked about the feasibility of the proposed data collection and test administration schedules, and to provide suggestions for the actual conduct of tests.

The set of instruments thus developed to measure program outcomes, institutional characteristics, educational program characteristics, and student characteristics are listed in Table II-6 (see Chapter II). The type of information provided by these instruments will be described in greater detail in the remainder of this chapter.

DESCRIPTIVE INFORMATION

Educational Program Characteristics

Measures of educational program characteristics were obtained from persons in a variety of roles. For the Facility Fiscal Questionnaire, fiscal officers at the institutions were asked to provide information on the level of Title I and other educational funding, and to specify the expenditure of these funds in terms of staff, training, equipment, and materials. These expenditures were detailed with respect to instructional programs in reading and mathematics.

Educational Program Administrators (EPAs) were asked to provide overall descriptions of the Title I and regular educational programs. Descriptions were obtained of participants and participant selection procedures; relative size and quality of programs in terms of staff, students, resources, and problems; program operations such as the length of school day, class offerings, and classroom structure; management characteristics of the administrator such as managerial style, leadership qualities and qualifications and attitudes of the EPA; perceptions of staff interactions and staff-student interactions, and the integration and coordination among educational programs within the institution. Finally, EPAs were asked to assess the relative importance of the various education programs and the importance of education itself vis-a-vis other institutional activities.

All teachers of the students to be tested in the impact study were asked to respond to the Teacher Questionnaire (TQA) describing the classes they taught, their instructional approach, their needs assessment procedures, the use of performance objectives, the degree of individualization of instruction, and their educational philosophy. Teachers were asked to describe the environment of their classrooms in terms of student involvement and cooperation. The management of educational programs was assessed from the teachers' perspective, as were teacher perceptions of administrative support for their teaching activities. Additionally, teachers provided descriptions of their own qualifications, experience, and background characteristics.

Because student improvement in basic skills was the major outcome variable, reading and mathematics teachers who were engaged in basic skills instruction were asked to provide descriptions of their classroom instruction in greater detail, utilizing the Title I Reading and Mathematics Program Characteristics Questionnaires (TQB, TQC). Measures were obtained for instructional methods utilized, degree of structure, degree of individualized instruction, resource utilization, availability and use of support services and personnel, and perceived support for basic skills instruction.

The Student Participation and Attendance Record (SPAR) was designed to obtain information on the length of student exposure to reading and mathematics instruction for a typical week prior to testing. Teachers rated individual students on their level of motivation and capacity to learn. At the completion of the testing cycle, institutions provided individual student entry and exit dates from educational programs and from the institution (SDS).

An observational/evaluational form was completed by SDC's data collection team on the second visit. Because teams were exposed to many educational programs it was felt that their observations of program operations and characteristics could serve as a validation of data provided by institutional personnel. Furthermore, these observers could provide assessments of difficult-to-measure aspects of program operations, e.g., degree of integration of Title I Programs with regular education programs, etc.

Institutional Characteristics

Correctional institutions frequently pursue multiple goals and objectives. In many cases, and particularly in adult institutions, education is but one of several programs, services, and activities operating simultaneously within an institution. For this reason, descriptions of the institutional context within which the educational programs operate were obtained from the institutional director (alternatively called a superintendent or warden), as well as from other respondents. Administrators provided descriptions of their institutions in terms of the characteristics of their residents (in terms of age, ethnicity, prior institutional histories, offender status, and length of stay, for example) and in terms of organizational characteristics of the institution: its size, crowdedness, levels of security, staffing patterns, and relative importance placed on various and sometimes competing activities.

Student Characteristics

Each student completed a questionnaire at the time of his or her first testing. It asked for information such as age, sex, ethnicity, educational history, offender status, history and length of institutionalization, participation in various programs, and perceptions of institutional goals.

STUDENT ACHIEVEMENT

Standardized Achievement Measures

This section of the report describes the selection of standardized achievement measures used in assessing the impact of Title I programs.

The Practical Achievement Scale (PAS), a measure of functional literacy, was previously developed by SDC and Pacific Consultants for the Sustaining Effects Study because a functional literacy instrument appropriate to compensatory education students with low levels of basic skills was not commercially available. This scale is domain-referenced and contains items derived from practical everyday experiences. Tasks such as reading a speedometer, reading package labels, and making simple monetary computations form part of the scale. These items

represent the viewpoint that functional literacy and practical achievement involve the reading and computational skills needed by students as they interact in the contemporary non-school-related world (Spencer et al., 1976). The availability of norms on this scale obtained from a recently sampled group of students participating in the Sustaining Effects Study allows performance comparisons of institutionalized students with students in the public school system.

The Comprehensive Tests of Basic Skills (CTBS), a standardized, norm-referenced test, was chosen for three reasons: (1) it was developed recently (1973) and reflects the growing concern of test makers that items be ethnically unbiased; (2) it enabled the performance of neglected or delinquent students to be compared to the performance of a nationally normed sample; and (3) it has been used in the SDC Study of the Sustaining Effects of Compensatory Education, enabling other comparisons between institutionalized and non-institutionalized pupils to be made.

Level 2 of Form S of the CTBS was chosen. The content of this test seemed appropriate for the students to be tested and comparison data on non-institutionalized Title I students (in grades 5 and 6) were available from SDC's Sustaining Effects Study.

Domain Referenced Tests (DRTs)

Domain-referenced testing is an outgrowth of recent developments in the field of educational evaluation. With the growing interest in individualized instruction and the attendant emphasis on diagnosis of deficiencies and prescription of remediations, it has become clear that standardized, norm-referenced tests are not adequate in assessing individual patterns of proficiency. At the same time, as interest has grown in comparing curricula, and in assessing the impact of educational programs, it has become clear that the standardized, norm-referenced test might not fairly represent the content of the curriculum or program being implemented.

The purpose of the domain-referenced test is to provide a clear, rigorous definition of an item form for a specific content area. The tester can then assess the performance of each student in terms of "measurements that are directly

interpretable in terms of specified performance standards" (Glaser and Nitko, 1971). Furthermore, it enables the curriculum designer and the teacher to assess the correspondence between the curriculum and the test. Millman (1974) has provided an explicit definition of a domain-referenced test:

> Any test consisting of a random or stratified random sample of items selected from a well-defined set or class of tasks (a domain) shall be referred to here as a DRT [domain-referenced test]. . . "Well defined" refers to an explicitly stated domain of items or tasks or to an item-generating procedure.
>
> The major advantage of the DRT is that it permits the user to make a special kind of CR [criterion-referenced] interpretation, namely an estimation of an examinee's domain score or level of functioning, defined as the percent of the population of items the examinee could answer correctly. . . . (p. 315)

After much consideration (detailed in Appendix A) it appeared that the basic skills in reading and mathematics were sufficiently well delimited so that explicit item-generation rules could be written. The resulting tests enabled us to estimate the proportion of each domain that had been learned by the students at each point of assessment. The DRTs provided much more specific information on the acquisition of basic skills than the standardized, norm-referenced tests.

The domain-referenced tests for use in Phase II were developed in three steps: domain identification and definition, domain specification, and item/test generation.

The identification and definition of the domains to be tested was a crucial step because the validity and usefulness of the tests is a function of the importance and meaningfulness of the domains to be tested. An extensive search of available curriculum and test materials in the basic skills area was undertaken, coupled with input from members of the Policy Advisory Group and the Research Advisory Group. This resulted in the selection of the basic skills domains for use in the Phase II study. The second phase involved the specification of the content to be tested, the item form(s) to be used, and the rules for generating individual items. The third step, test generation, involved

the sampling of items from domains and assembling the items into test forms. This phase included the overall testing strategy, the design of test instruments, and the characteristics (number of items, length of time, etc.) of an individual test.

In Appendix A, each of the above steps is described in detail. In the following section of this chapter, a brief description of the DRTs, and a discussion of their technical qualities is presented.

The Domain Referenced Test of Arithmetic Basic Skills

The DRT for arithmetic basic skills was constructed to include the four basic operations of addition, subtraction, multiplication, and division, and four types of item content: 'fact', integers, decimals, and fractions. The test was designed such that each operation was performed on each of the item content types. The test was organized into four booklets, by type of operation. Thus, each booklet contained items of one type of operation but of all four types of content. No reading was required in responding to these items.

The arithmetic domain-referenced tests were constructed according to a specified hierarchical structure, both within a test and across tests. The addition, subtraction, multiplication, and division tests had four discrete parts. The first part consisted of eight arithmetic 'fact' items--operations on single digit integers--typically found on 'numbers tables' posted on elementary school classroom walls. The second part consisted of items involving integer numbers, at least one of which has more than one digit. The third part consisted of operations on decimals, and the fourth part consisted of operations on fractions. The specific delimitations of the content are given in Appendix A.

The logic of the order of these content types is as follows. Mastery of number 'facts' is usually considered prerequisite to more complex operations, including larger integers, decimals and fractions. Thus they were tested first. In order to perform operations on higher valued integers, an additional manipulation is usually required: carrying in addition and multiplication, borrowing in subtraction, and computing the remainder in division. The items with decimal

numbers involve knowledge of number facts, the above manipulations and knowledge of the proper placement of the decimal point. Correct manipulations of decimals logically presumes mastery of fact and integer items. Items involving fractions, while not building directly on decimal manipulations, have the same prerequisites as decimal items plus additional manipulations. All operations on fractions require understanding of the concepts of numerator and denominator. The addition and subtraction of fractions require an understanding of how to compute and use the common denominator. The division of fractions requires an understanding of the reciprocal. Because fraction items require more additional concepts and operations than decimals, they were placed last in the ordering of content types.

A similar, hierarchical structure was presumed for the operation types: in order to perform long division, a student must be able to multiply and subtract. To multiply, a student must be able to add. If addition is considered as prerequisite to subtraction (the ordering usually found in elementary arithmetic curricula) then the ordering postulated is addition, subtraction, multiplication, and division.

The hierarchical structure of operation and content types was investigated by calculating the mean proportion correct for each type. The means of the operations are presented in Table III-1. These means provide evidence of the hierarchical ordering among operations. Students correctly answered fewer later-ordered items. The means for content types (presented in Table III-1) also confirm the expected order of difficulty. In general, students correctly answered most of the fact items, fewer of the integer items, fewer still of the decimal items, and very few of the fraction items. This order is preserved for every test, except multiplication. Common fractions are easier to multiply than to add, subtract, or divide. One simply multiplies numerators and denominators, respectively. More complex manipulations are required for addition, subtraction, and division.

Table III-1. Mean Proportion Correct on Arithmetic Subtests at Wave 1

Operation Type	Content Type					Operation Total
	n	Facts**	Integers	Decimals	Fractions	
Addition	1149*	.98	.91	.61	.36	.78
Subtraction	1147*	.96	.80	.54	.28	.71
Multiplication	1130*	.87	.55	.28	.32	.59
Division	1065*	.85	.40	.41	.13	.53

Because the hierarchy among arithmetic tests was verified,*** a sum score over all tests was determined to be a valid indicator of each student's acquisition of arithmetic skills. Thus in all analyses of student achievement the sum score of the four arithmetic tests was used rather than the subtest scores.

A fifth test booklet was produced in which basic skills operations were applied in descriptions of practical settings. This test had 12 story problems. The content ranged from making change to computing simple interest. The score on this test was considered separately from the other arithmetic DRTs in assessing outcomes. Unlike the other arithmetic tests, these problems required reading skills.

The specifications of this test of practical mathematics are also given in Appendix A. Information on the reliability of this test is given in this chapter.

---

*Title I and non-Title I students from only the 30 randomly selected institutions are included in this analysis.

**Operations on single digit numbers.

***See Appendix B for further discussion of analysis results verifying the hierarchical structure among tests and subtests. These analyses examined performances on earlier subtests for students who had mastered later subtests, and also present correlations among tests and subtests.

The Domain Referenced Tests of Language Arts Basic Skills

The domains included in the language arts DRTs were word recognition, vocabulary, and literal comprehension. The word recognition and vocabulary tests each yielded one score. The literal comprehension test consisted of eight passages with certain words deleted. The student had to find the most appropriate word to insert from a list of words provided at the point of deletion. The eight passages were graded on a scale of reading difficulty. (See Appendix A for further discussion of reading difficulty.) Passages from a variety of content areas were included, selected from materials found in grades 2 through 10. Appendix A gives the details of the construction of this test.

The hierarchical structure was investigated by calculating the mean proportion correct for the eight passages. The means of the eight passages in Table III-2 display a hierarchical order corresponding to the order of the presumed difficulty of the passages. Consequently, it appears that the sum score across all passages captures the salient information on this test. The word recognition and vocabulary tests were not constructed of subparts; discrete items were ordered in ascending difficulty. Although good performance on the literal comprehension test clearly presumes skill at word recognition and knowledge of

Table III-2. Mean Proportion Correct on Literal Comprehension Passages at Wave 1*

Passages	P
1	.87
2	.72
3	.72
4	.68
5	.60
6	.63
7	.54
8	.50

*Title I and non-Title I students from only the 30 randomly selected institutions are included in this analysis.

vocabulary, this test was thought to include a wide variety of skills not tapped by the previous two tests. Combining information on the three tests would tend to dilute information given by the literal comprehension test. Thus total scores on each of the three tests were used as indices of language arts skills in all analyses of student achievement.

Form-to-Form Difficulty and Reliability of the Domain-Referenced Test

The domain-referenced tests were administered four times at six-week intervals. To minimize the effect of practice, four parallel forms of each test were constructed. At the first test administration, forms were randomly assigned to students so that each form was assigned to one-fourth of the students in the sample. A Latin square design was used to assign forms to students on subsequent testings to ensure that students had a different form at each testing.*

To assess the magnitude of differences across forms, analyses-of-variance were conducted comparing scores on the four forms during the first test visit. Table III-3 gives the mean proportion correct for each form of a test and the significance of the F ratio for the analysis-of-variance of each test. Different forms of a test showed strikingly comparable levels of difficulty. The mean proportions correct were nearly identical across all forms of the arithmetic tests. Among language arts tests, paragraph comprehension and vocabulary showed differences across forms. Examination of the means for those tests revealed that Form 2 was significantly easier than the other forms in paragraph comprehension. In vocabulary, Form 2 was significantly harder than the other forms. Thus, scores in paragraph comprehension and vocabulary were adjusted for the form

*Due to an oversight in the assignment procedures, many students received the same form on the first and third test administrations. To test whether a practice effect influenced achievement, scores at all four test administrations were intercorrelated, and means across the four testings were compared. The intercorrelations showed little evidence of this effect; the correlations between first and third tests were no higher, and sometimes lower, than correlations between other test administrations. Comparisons of mean scores across four waves of testing showed evidence of a practice effect of the same form only for paragraph comprehension. For this test, mean scores on the first and third test administrations were more alike than scores on other testings.

CO

Table III-3. Differences Across Parallel Forms of the Domain-Referenced Tests at Wave 1*

Test	Mean Proportion Correct for Form				Significance of F Ratio
	1	2	3	4	
Addition	.79	.78	.79	.78	NS
Subtraction	.73	.72	.72	.72	NS
Multiplication	.59	.56	.58	.57	NS
Division	.56	.54	.53	.53	NS
Practical Mathematics	.29	.30	.30	.29	NS
Word Recognition	.94	.95	.94	.95	NS
Vocabulary	.66	.63	.67	.69	.001
Paragraph Comprehension	.67	.76	.68	.69	.001

differences in subsequent analyses. Because analyses of the other tests showed no statistically significant differences between forms, those scores were not adjusted.

Because the domain-referenced tests were administered more than once, information on reliability of the test scores was assessed in several ways. First, a measure of internal consistency, Cronbach's coefficient alpha (α) was computed. This coefficient is an index of the homogeneity of items on a test. Second, a measure of consistency over time, the correlation between scores at different test administrations, was computed. It should be noted that this correlation is not a simple test-retest correlation which is usually interpreted to be influenced only by random fluctuations in the condition of the student or of the testing

*All students who had taken a test were included.

environment. This correlation includes the influence of instruction on the knowledge of the content of the domain-referenced tests. Because the exposure to instruction differed across students and institutions, one would not expect a perfect correlation over the six-week intervals even if the test correlated perfectly on two occasions close-in-time. Both measures of reliability are given in Table III-4. The measures of internal consistency are fairly high for all tests. Each test seems to be homogeneous in content. Thus, although the arithmetic tests each consisted of items on facts, integers, decimals, and fractions, the central arithmetic operation, i.e., addition, produced consistency among these different types of items.

Table III-4. Reliability of Domain-Referenced Tests

Test	No. Items	Reliability Coefficient						
		Internal Consistency (α)*	Correlations Between Tests at Waves**					
			1,2	1,3	1,4	2,3	2,4	3,4
Addition	28	.86	.66	.75	.70	.76	.73	.81
Subtraction	25	.90	.74	.73	.73	.77	.76	.78
Multiplication	24	.89	.78	.75	.73	.80	.80	.76
Division	24	.93	.84	.78	.75	.80	.80	.82
Practical Mathematics	12	.78	.73	.69	.68	.74	.75	.78
Word Recognition	16	.90	.56	.64	.49	.75	.63	.65
Vocabulary***	24	.81	.60	.67	.56	.65	.64	.57
Reading Comprehension***	75	.97	.65	.74	.69	.71	.74	.75

*Coefficient alpha (α) was calculated for test scores at the first wave of testing.

**The coefficients in this table were computed on a minimum of 280 cases and a maximum of 840 cases.

***Because form-to-form differences were found for reading comprehension and vocabulary, coefficient alpha (α) was computed for a single form and across forms. All analyses produced nearly identical results.

The intercorrelations among test administrations shown in Table III-4 are lower than the measures of internal consistency, which is to be expected. These figures show that scores on most of the tests were fairly stable over time. The lower correlations for practical mathematics and word recognition may be partially explained by the restriction in range of scores for any one test administration. Many students performed at or near the maximum score on word recognition, showing evidence of a ceiling effect. On practical mathematics, many students performed at or near the minimum score, showing evidence of a floor effect. Because the range in scores on these tests was less than that of the other tests, somewhat lower correlations might be expected.

Generally, the correlations for tests separated by six weeks are higher than those separated by twelve weeks, which are higher than those separated by 18 weeks. This pattern is expected in repeated measures data of this nature (Bock, 1963 and 1975).

In summary, the DRTs were first examined to assess the degree to which the subtests were hierarchically ordered. The results indicated that total scores in a basic skills area would provide complete information on student performance in those areas. Second, form-to-form differences were examined. Results of these comparisons showed that different forms of the DRTs were equivalent. Third, an examination of internal consistency and test-retest reliability showed that the test scores were reliable indicators of student performance.

It should be emphasized that these tests were generated by the application of specific item construction rules, as demonstrated in Appendix A. They were <u>not</u> generated by the usual item analysis procedures which eliminate items based on their difficulty levels or correlations with total test scores. The technical quality of these tests indicates that the exclusive use of item construction rules can result in useful, reliable instruments.

STUDENT ATTITUDES

Development of Student Attitude Measures

Based on advisory group recommendations and previous research done in juvenile corrections, several attitudinal areas were identified for inclusion in the Phase II evaluation: self-esteem, locus of control, attitudes toward school and basic skills tasks, and perceptions of the institutional environment. The task of designing instruments to measure these areas proved to be difficult, however. That task involved reviewing other instruments and the research literature associated with them and modifying the instruments to account for the institutional settings and populations of youths that are the focus of this study. These steps are described in detail below.

Generally speaking, there are few well-developed measures of student attitudes. Instruments are often multi-dimensional in content (at the expense of measuring a few constructs well) and the use of scale scores varies widely from study to study--subscales are treated as unique in some studies and are added together in others. Furthermore, the practice of tailoring questions to particular populations and circumstances raises questions about the generalizability of construct interpretations. The most important criteria used in the selection of instruments and scales for this part of the Phase II study were the ease of administration and scoring, the psychometric properties of scales, the item format, and the item content.

Time constraints for both students and data collectors and the literacy level of the students being assessed led us to look for straightforward, paper-and-pencil instruments in which the items were clearly stated (all items were to be read aloud to the students as well as presented in written form) and the required student responses would be of the true-or-false variety. Projective instruments such as the TAT, the Rorschach and the Draw-a-Person test were thus eliminated.

The need for low reading-level items and simple response formats was complicated by the fact that students participating in the evaluation were expected to be older than populations studied in other compensatory education evaluations.

Thus, although the reading performance of the population was expected to be on a primary-grade level, their experiences were expected to be on a more adult level than would be true of their functional-level counterparts in public schools. This meant that item content had to be reviewed so as not to be perceived with disdain by the 'streetwise' respondents.

The validity and reliability of existing instruments were reviewed and scales lacking acceptable psychometric properties were eliminated from further consideration.

Another important concern in the selection of items and administration of the test was that of response bias. Students in correctional institutions may feel that reporting favorable attitudes would lead to early release. Attempts to minimize this bias were made by selecting carefully worded items and by assuring students that their responses were anonymous. The wording of the items was chosen to reduce the tendency to respond in a uniform way regardless of the item content.

The concept of self-esteem has been noted in previous research to include such areas as peer acceptance, self-acceptance, self-derogation, personal ability, and family-related attitudes. (Since family-related attitudes could not realistically be expected to impact on Title I objectives, this area was eliminated from consideration and resulted in the deletion of many items in published instruments.) A modified version of Coopersmith's Self-Esteem Inventory (as modified for use in correctional institutions by Bennett, et al., 1975) was used as a measure of self-esteem. The instrument was modified further by eliminating items pertaining to family acceptance and the self-acceptance of physical bodily features. Additionally, items that reflected absolute statements (I'm always unhappy) were modified to reflect tendencies (I'm usually unhappy).

Items measuring locus of control (the perception of one's ability or inability to influence outcomes) were selected from A Locus of Control Scale for Children (Nowicki and Strickland, 1973), while items measuring attitudes toward school, learning, and basic skills were selected from the Survey of School Attitudes (Harcourt, Brace, and Jovanovich). All items selected on these three dimensions were subjected to a pretest on 278 students residing in four correctional

institutions. Results from a factor analysis of this pretest data served as a basis for further elimination and clarification of items. The dimensions to be included in the assessment of student attitudes toward educational programs in institutions were selected on the basis of hypothesized relationships to program impacts and to previous demonstrations that correctional environments varied in these dimensions (Wenk and Moos, 1972). Table III-5 presents these dimensions and a brief definition of each. Items assessing these dimensions were adapted from the Classroom Environment Scale (Trickett and Moos, 1971).

Table III-5. Dimensions of Student Perceptions of the Educational Programs Within the Institution

Dimension Label	Content
Involvement	Extent to which students participate in the daily functioning of educational programs.
Expressiveness	Extent to which the education programs encourage the open expression of feelings by residents and staff.
Affiliation	Extent to which students work with and come to know each other within the educational environment.
Rule Clarity	Extent to which rules within the school situation are explicitly stated and clearly understood.
Staff Support	Extent to which staff express personal interest in the student's academic achievement and personal life.
Practical Orientation	Extent to which institutional programs are oriented to the practical concerns of students once they are released.
Classroom Order and Organization	Extent to which classroom emphasis is on the maintenance of order and the degree to which class activities are organized.
Institutional Support for Education	Extent to which academic educational activities and goals are pursued rather than other institutional goals.

Because of the large number of items in the total survey of attitudes, two test booklets were constructed: Student Attitude Measures I and II. Each booklet contained a mixture of items from all dimensions. One of these booklets was administered on each of the two days of testing on the first visit. This procedure was repeated on the third visit.

All hypothesized scales were subjected to a factor analysis to investigate their underlying structure and dimensionality. The results of this analysis are presented in Table III-6. Obtained self-attitude dimensions were self-esteem, locus of control, sociability, and expectations for post-release success or stigmatization. The dimension 'attitudes toward school' split into two dimensions: school involvement, and liking of school. Attitudes toward reading tasks and math tasks formed two separate dimensions as anticipated. Structures emerging in the psychosocial perceptions of institutional environments were peer affiliation, practical orientation, and teacher support. The internal consistency reliability of these scales ranged from .48 to .80. A detailed description of the results of the factor analysis, the scale reliabilities, and inter-scale correlations are presented in Appendix C.

Table III-6. Student Attitude Scales

Scales and Items

SELF-ESTEEM:

There are a lot of good things about myself.
Most of the time I'm proud.
I usually know what I want to do.
I usually feel good about myself.
I'm not very proud of myself.
I can do most things as well as other people.

LOCUS OF CONTROL:

I only work when my teachers ask me to.
It's almost useless to try to learn in school because I can't do the work.
Being sent here has ruined my whole life.
Many times I feel that I have little say over things that happen to me.
When bad things are going to happen, they just happen no matter what I try to do to stop them.
Teachers here order me around a lot in class.
When I do something wrong, there is very little I can do to make it right.
One of the best ways to handle most problems is just not to think about them.
When people are mean to me, it is usually for no reason at all.
No matter how hard I try, no one notices my efforts.
I am often blamed for things that aren't my fault.
It usually doesn't pay to try hard because things never turn out right anyway.

SOCIABILITY:

I have a lot of friends.
I'm easy to like.
Most people like to be around me.
I'm a lot of fun to be with.

PEER AFFILIATION:

Students here get along well with other students.
I get along with other students in my classes.
Students of different races get along with each other here.

SCHOOL INVOLVEMENT:

I work hard in class.
I try hard to do better in school.
I'm doing the best work I can in school.

Table III-6. Student Attitude Scales (Cont'd)

LIKING OF SCHOOL:

I would like to stop taking some of my classes.
My classes are usually dull and boring.
I am usually unhappy when I am in school.
Most days I do not want to go to school.
I often get upset in school.

ATTITUDES TOWARD MATH:

I don't like multiplying numbers.
I like doing problems with fractions.
I like dividing numbers.

ATTITUDES TOWARD READING:

I like reading stories.
I like studying the meaning of words.
I like reading newspapers and magazines.

TEACHER SUPPORT:

Most of my teachers here are fair.
Teachers here go out of their way to help me.
Teachers here care what happens to me when I leave here.
If my teacher does something I don't like, I can tell him/her about it.
I can talk to teachers about things that bother me.
My teachers tell me when I am doing well.
Most people like to help me out when I need help.

PRACTICAL ORIENTATION OF LEARNING:

I am learning a lot in my classes.
I am learning things I'll need to know when I leave here.
What I have learned in school will help me when I leave here.
I have learned things in school here that will hlep me when I leave here.

EXPECTATIONS FOR RELEASE SUCCESS AND STIMATIZATION:

It will be hard for me to stay out of trouble with the law now that I have been in a place like this.
I think people will give me a fair chance when I leave here as long as I stay out of trouble.
I think that having been here will hurt my chances for getting a good job when I leave here.
Most people on the outside don't give someone who has been in a place like this a fair chance.

CHAPTER IV. CHARACTERISTICS OF STUDENTS, PROGRAMS, AND INSTITUTIONS

PRECIS OF CHAPTER IV

This chapter presents summary data on many characteristics of students, educational programs and institutions. Among student characteristics, it was found that institutionalized students are on the average 17 years of age (although the standard deviation is large), and that Title I students are, on the average, one year and seven months younger than non-Title I students at the randomly selected institutions. This finding indicates a need to adjust for age differences in any analyses of the performance differences between Title I and non-Title I students. On the average, students had been confined in institutional settings on more than two occasions. Title I students had significantly greater numbers of prior institutionalizations than did regular education students. Students reported their last grade attended prior to institutionalization to be the ninth grade on the average. About 75 percent of students reported property or person-related offenses to be the reason for institutionalization.

Many of the institutional characteristics reported in this chapter resulted in distributions that were very skewed and had large standard deviations. The median length of stay was reported to be 7.5 months while the mean and standard deviations were 13.7 and 15.2 months, respectively. The size of the institutions was also very skewed, with a mean of 407 and a median of 164 residents. About one-third of the institutions reported that they were overcrowded.

Despite the recent emphasis on community-based treatment and on release programs, most institutions offered all academic and vocational instruction on site. One-third of the institutions reported having programs that were not state accredited; only one-half of the institutions received any information from community schools on institutionalized students; and few institutions said that they returned information to community schools on any of their released students, all of which suggests a lack of coordination between institutional and community schools.

Although there is some question of the reliability of expenditure data obtained, it was found that the average Title I student receives supplementary services worth about one-half of the cost of the services provided to all students by the regular education program. However, because much of the supplement is targeted to basic skills instruction, the Title I student receives supplementary reading instruction worth 1.5 times the cost of the reading instruction provided to all students through regular education. In mathematics the supplement is worth nearly twice the amount of the regular education program. These cost data compare poorly to the data on instructional time, which, while incomplete, seem to indicate that students who are to receive Title I supplementary services in both reading and mathematics receive (or accept) only a small portion of the services which are offered.

Examination of the instructional characteristics of educational programs revealed that, in general, planning by compensatory education teachers was sensitive to the instructional needs of students in compensatory education. A comparison of teachers of compensatory education of confined youth to similar teachers in public school settings revealed that institutional teachers place more emphasis on individualized instruction and less emphasis on small group instruction than their public school counterparts.

Assessments of procedures used to assign students to Title I instruction revealed that nearly half of the institutions assigned all students or all eligible students to Title I programs. The remaining institutions used a combination of test scores, staff recommendations, and student willingness to participate in order to identify Title I recipients. The seemingly unsystematic way in which students are chosen for Title I treatments has implications for the evaluation of the influence of programs on achievement outcomes.

Analysis of the somewhat sketchy data on amount of student exposure to instruction suggests that other institutional activities and problems with student attendance interfere with the instructional program and may result in limited exposure to instruction.

In Chapter V, these data on student, institutional, and educational program characteristics are related to the achievement scores and attitude measures obtained on the students.

CHAPTER IV

STUDENT CHARACTERISTICS

Characteristics of students have been shown in various research efforts to relate to outcome changes in student achievement and affect. In this section student characteristics are described for the neglected or delinquent youth population as a whole and, separately, by educational program--Title I or regular education. These descriptions show the variety of students receiving educational services in institutions for neglected or delinquent youth.

Information on characteristics of students was obtained from a questionnaire administered to every student present for testing. This questionnaire asked information on age, sex, ethnic background, last grade attended before institutionalization, number of schools attended since first grade, reason for institutionalization, time remaining to be spent in the institution, and number of times institutionalized. To obtain valid responses on the questionnaire, students were assured that all responses would be anonymous.

Tables IV-1 through IV-3 summarize student characteristics for students receiving Title I instruction, those receiving only regular education instruction, and for all students without distinguishing between educational programs. These tables present data from the randomly sampled institutions, weighted to account for the different probabilities of student selection within institutions. Chapter II provides more detail on the sampling design and the weighting factors. Appendix D presents summary statistics of student characteristics for each of the randomly selected institutions.

The average age of neglected or delinquent students within the sample was about 17.3 years. Students in institutions for the neglected were youngest, with ages ranging from 10 to 17, averaging about 15 years. Students in delinquent

Table IV-1. Distribution of Student Background Characteristics by Educational Program
(Data are from 30 randomly selected institutions)

Characteristic	n	Mean	SD
Age			
Title I	1291	16.6	2.5
Regular Education	671	18.2	3.6
All students	2182*	17.3	3.1
Last grade attended			
Title I	1308	9.0	1.5
Regular Education	667	9.6	1.6
All students	2199	9.2	1.6
Number of schools attended			
Title I	1272	5.3	2.9
Regular Education	670	5.4	3.0
All students	2162	5.4	3.0
Number of times institutionalized			
Title I	1208	2.6	2.9
Regular Education	619	2.3	2.6
All students	2033	2.5	2.8
Months left in institution			
Title I	1210	7.5	11.3
Regular Education	640	8.5	10.4
All students	2060	8.2	11.9
Sex	n	% Male	% Female
Title I	1330	85	15
Regular Education	688	89	11
All students	2247	87	13

*Program status for some students could not be ascertained; those students are included in "all students".

Table IV-2. Percent of Students in Each Category of Reason for Institutionalization and Educational Program (Data from randomly sampled institutions)

Program	n	Reason for Institutionalization				
		Neglected or Dependent	Truancy Runaway Incorrigible	Burglary, Larceny Auto theft, Arson Fraud, Embezzlement	Assault, Battery Robbery, Sexual Assault, Homicide Manslaughter, Rape	Drug Offense
Title I	1128	1.6	15.6	48.3	28.7	5.9
Regular Education	614	3.8	12.2	50.1	25.5	8.4
All Students*	1946	2.3	14.4	48.9	27.6	6.8

*Program status for some students could not be ascertained; those students are included in "all students".

Percentages may not add to 100 due to rounding.

Table IV-3. Percent of Students in Each Category of Ethnic Background by Educational Program (Data from randomly sampled institutions)

Program	n	American Indian or Alaskan Native	Asian or Pacific Islander	Black	White	Hispanic
Title I	1303	7.7	0.2	42.3	40.1	9.7
Regular Education	671	7.9	0.6	35.0	49.8	6.7
All Students*	2194	7.8	0.3	39.8	43.4	8.7

*Program status for some students could not be ascertained; those students are included in "all students".

Percentages may not add to 100 due to rounding.

institutions ranged from 13 to 21, averaging about 16.5 years. Students in adult institutions were oldest, averaging about 20 years. The range in average age over institutions was large, from about 13 to over 21. This large variation in age suggests that educational programs and expectations about performance should be adapted to the ages of the students involved. A striking finding is the statistically significant ($p = .001$) difference of 1.6 years between the ages of students in regular education programs only and students in Title I programs. Because older students can be expected in general to perform better than younger students, all analyses of outcomes comparing Title I to regular education only students will contain an adjustment for age.

Most of the neglected or delinquent students were male. In general, the population of students at every adult institution, and at half of the delinquent institutions, were entirely male; most of the remaining facilities housed both males and females, with males generally in the majority.

Students reported that, on the average, they had last attended ninth grade before coming to the institution, but facilities differed widely on this variable. The .6 grade difference between Title I students and regular education students was statistically significant at $p = .001$, and was found to be due to the age difference between the groups. Students had, on the average, attended five different schools before coming to the institution. Although facilities differed on this variable, no difference between Title I and regular education students was found.

Students had been institutionalized on prior occasions an average of two and one-half times. Title I students had been previously institutionalized more often than regular students, which proved statistically significant at $p = .001$. Students anticipated an average of about eight months left at the institution at the time of the testing. This length of time varied considerably across institutions, ranging from a low of 1.6 months at a delinquent facility to 26.6 months at an adult facility. Title I and regular education students anticipated about the same number of months left in the institution.

In Table IV-3 the percentage of students in each category of ethnic background is given. Black and white students made up the bulk of the neglected or delinquent sample. The proportion of white students in programs of regular education only was significantly higher than the proportion of black students in such regular education programs. The percentages for each institution show considerable differences in ethnic composition across institutions.

Half of the most serious offenses that students reported as the reasons for their institutionalization were property-related (burglary, larceny, auto theft, arson, fraud, embezzlement). Just over one-fourth were person-related offenses (assault, battery, robbery, sexual assault, homicide, manslaughter, rape). That only about 7 percent of respondents reported drug-related offenses as their most serious offense may substantiate the pervasiveness of the movement towards de-facto decriminalization of drug offenses. That 14 percent of students in institutions for delinquent youth report status offenses as their most serious offense indicates that the deinstitutionalization of status offenders is not yet complete. The relative proportions of offenses were similar across types of institutions, with the exception that adult institutions housed a higher proportion of person-related offenders than did neglected or delinquent institutions. The distribution of offenses was similar across Title I and regular education students.

DESCRIPTIONS OF INSTITUTIONS

Table IV-4 gives the medians, the means and the standard deviations for the responses to certain descriptive questions obtained from the facility directors at the 30 randomly sampled institutions. Many of these variables have large standard deviations relative to their means and have very skewed distributions. For example, the typical length of stay (in months) has an average value of 13.7, a median of 7.5 and a mode of 8, meaning that the most frequent stay is 8 months. One-half of the students stay 7.5 months or more and, due to a few extremely long-duration times, the arithmetic average is 13.7 months. (The values range from 1 to 60 months.) Size of the institution proved to have a similarly skewed distribution, with an average of 407 and a median of 164 residents.

Table IV-4. Descriptive Information on 30 Randomly Sampled State-Operated Institutions for the Neglected or Delinquent (Facility Director Responses)

Item	Median	Mean	s.d.	n
Typical length of stay of resident (in months)	7.5	13.7	15.2	30
Maximum number of residents facility is designed for	182	344	419	30
Number of residents housed at facility.	164	407	613	30
Percentage of residents kept under medium or maximum security	8.0	32.9	40.7	30
Percentage of turnover annually in:				
custody	3.0	8.4	12.0	27
cottage staff	0.0	10.0	17.0	27
non-education treatment staff	0.0	8.0	12.5	27
education staff	2.0	7.6	11.7	27
Percentage of custody staff who have a high school diploma	96	74.8	39.0	30
Percentage of cottage staff who have a high school diploma	50	49.6	48.7	27
Percentage of facility's staff which is white, not of Hispanic Origin	82	76.9	18.7	27
Percentage of facility's staff which is female.	21	29.7	26.9	27
Percentage of residents who participate (in a typical month) in:				
Work release programs	0.5	6.3	11.2	27
School release programs	0.0	4.0	8.7	27
Family release programs	0.0	12.9	25.3	27

Perhaps one of the more interesting findings about the characteristics of the sites is the overcrowding in these facilities. When the means are examined, there appears to be a sizable amount of overcrowding, suggesting that there are a few sites with serious problems in this area. Nine of the 30 randomly selected institutions reported that they had more residents than they were designed to hold. Of these, four small facilities (fewer than 250 residents each) had from 1 percent to 16 percent residents above capacity. Another facility reported that it housed 30 residents even though it was not designed as a residential facility. The other four facilities are very large (1110 to 2600 actual residents) and have between 35 percent and 94 percent more residents than they were designed to accommodate.

With respect to other institutional characteristics which could influence the educational programs, it was reported that there was little turnover in the educational staff and that most of the custody staff was educated at the high school level. There is little participation in 'release' programs, thus, most of the educational programs serving residents are offered by the institutions themselves. Only institutions serving neglected students used community schools for part of their program.

CHARACTERISTICS OF EDUCATIONAL PROGRAMS

Residents in Programs

Table IV-5 summarizes the data from the randomly sampled institutions concerning the size of educational programs and the percentage of residents participating in programs. The typical institution provides academic education to a substantially larger proportion of its residents than it provides with vocational education. The procedures by which students are assigned to programs of regular education--plus Title I supplementary instruction--are discussed in a later section of this chapter.

Table IV-5. Participation of Residents in Educational Programs
(Data are from 30 randomly sampled institutions.)

	Median	Mean	s.d.	n
Numbers of Residents Participating in				
All education programs on a typical day	99	134.0	124.3	30
Title I instructional program on a typical day	30	52.5	55.5	30
Total Number of Residents Served in a year in				
All educational programs	350	429.2	384.4	30
Title I instructional program	100	186.5	183.2	30
Percentage of residents participating in				
Academic education	85.5	69.3	34.1	27
Vocational education	31.5	41.1	34.1	27
Delay (in weeks) from time of entry to institution until a student enters				
Regular education program	2	2.7	2.55	30
Title I education program	1	2.8	3.81	30

Table IV-6 presents data on some other aspects of the educational program. Education Program Administrators at the 30 randomly selected institutions also reported that 27 percent of the institutions had no state accredited educational programs. They reported that only one-half of the institutions receive educational histories on all their students from the schools the students previously attended. Thirteen percent of the institutions never send information on the educational experiences of any of their students to schools outside the institution, and 27 percent never receive feedback that a former student has enrolled in school after release. These data reveal a lack of coordination between regular public schools and institutional schools. This lack of coordination may contribute to a low rate of return to regular schools among the formerly institutionalized students. One of the major objectives of a follow-up study conducted as part of this evaluation has been to assess the effects of pre-release and transitional services provided to institutionalized youth who wish to reenter regular schools. This work is reported in a companion volume to the present document, Post-Release Experiences of Students From State Correctional Institutions, Volume 3 in this series, 1979.

Operating Costs of Programs

The person responsible for fiscal operations at each institution was requested to provide information about the costs of educational programs. The focus was on academic education and the Title I supplements to such programs. Table IV-7 shows the summary of the data obtained on expenditures by program, while Table IV-8 gives a summary of the expenditures on a per-pupil basis. Thus, the cost of all educational activities was about $300,000.00, on the average, while the cost of the Title I supplement was about $71,000.00, on the average. The average student received about $2,980.00 worth of services, while those students who were in Title I instruction received about $1,538.00 worth of supplementary services. Stated another way, the average per-student cost of services received by Title I students was approximately one and one-half times the cost of the services received by students only in regular education programs.

Obtaining reliable data on estimated educational expenditures from correctional institutions has proven to be a difficult task during both Phase I and Phase II of the evaluation. The major sources of difficulty were that the system of

Table IV-6. Other Characteristics of Educational Programs
(Data from 30 randomly sampled institutions.)

	Median	Mean	s.d.	n
Percentage of Teachers who are White, not of Hispanic Origin	76	7[illegible].4	22.1	30
Percentage of Teachers who are female	25	30.5	25.5	30
Percentage of Teachers under age 35	50	50.7	29.0	30
Monthly days of absence per				
regular education staff member	1	1.3	.98	30
Title I staff member	1	1.0	.90	30
Educational Program Administrator (EPA) Experience (in years)				
In current position	3	3.8	4.4	30
Total Teaching	8	9.2	5.7	30
Teaching in correctional settings	4	5.9	6.1	30
Total administrative	4	5.8	4.6	30
Administrative in correctional settings	4	5.0	4.3	30
EPA perceptions of degree of integration of the Title I program with regular education program (1 = integrated - 3 = independent)	1	1.4	0.49	30
EPA perceptions of degree of interference of other institutional activities with educational program (1 = never - 4 = always)	2	2.3	0.61	30

Table IV-7. Summary of Data from Fiscal Officers at 30 Randomly Sampled Institutions

Variable	Mean	s.d.	n
Estimated total educational expenditures 1977-78			
Title I N or D funds	71,015.00	63,287.50	25
Other education funds*	300,164.00	249,385.50	25
Estimated Expenditures on Reading Program			
Title I N or D funds	29,189.00	23,861.00	24
Other education funds	48,847.00	54.674.00	23
Extimated Expenditures on Math Program			
Title I N or D funds	20,920,50	20,484.50	24
Other education funds	34,578.50	48,321.50	24
Estimated Expenditures on Intergroup Relations Programs			
Title I N or D funds	4,726.00	11,834.50	25
Other education funds	63,791.50	150,930.00	26
Estimated Expenditures on Other Educational Programs			
Title I N or D funds	6,830.00	13,871.50	23
Other education funds	**	**	

*Other educational programs funded by sources other than Title I N or D funds would include vocational education, physical education, science, etc. These expenditures account for the difference between the total educational expenditures and the parts shown in the remainder of the table. The data reported by expenditure category did not always add up to the total expenditure for the Title I program, thus the averages given also do not add up to the average total.

**Data not obtained

Table IV-8. Summary of Data on Costs Per Student at 30 Randomly Sampled Institutions

Variable	Median	Mean	s.d.	n
Estimated total education expenditures 1977-78.				
Title I funds per Title I student	1,162.00	1,538.30	1,113.80	24
Other funds per student	2,322.60	2,980.30	2,470.90	25
Estimated expenditures on Reading Program 1977-78				
Title I funds per Title I student	622.20	781.40	682.80	23
Other funds per student	401.10	583.50	766.50	23
Estimated expenditures on Mathematics Program 1977-78				
Title I funds per Title I student	476.80	510.20	454.20	23
Other funds per student	246.00	411.70	592.10	24

The Title I funds were divided by the typical daily enrollment in Title I programs. Other funds were divided by the typical daily enrollment in all educational programs, including Title I. The total funding for a Title I student is presumed to be the sum of the Title I funding per Title I student plus the other funding per student. Thus, while the average, non-Title I, per-student expenditure for reading programs was $583.50, Title I expenditures added a supplement of $781.40 per student, on the average. The large difference between the sum of the non-Title I per-student expenditures in mathematics and reading programs, and the total expenditures is accounted for by large expenditures for vocational education, physical education and general education in which reading and mathematics instruction are not differentiated (e.g., GED preparation).

record keeping and the categories of expenditures used in maintaining records varied widely among administrative agencies and among states. Within some administrative agencies or states, record keeping was centralized, and categories of expenditures were aggregated above the institutional level. Thus, the first source of difficulty in these situations was in obtaining separate expenditures for a particular institution from the agency or state. Secondly, it was difficult to ascertain the extent to which variation in estimated expenditures was due to discrepant categories of expenditures and the subsequent inability of fiscal administrators to break down these expenditures into categories of analytic interest.

In addition, the costs are likely to vary by type of institution, level of security, and region of the country in ways that our data are not accurate enough to ascertain. Despite the problems, the data reported here should provide a useful overall estimate of expenditures on education.

The data summarized in Tables IV-7 and IV-8 include only those institutions returning usable information. Enumeration of more specific problems with expenditure data revealed that four facilities did not return questionnaires. One institution reported that its Title I program was only for summer school. A few other institutions could not provide data on the costs for reading and mathematics instruction separately, as these were subsumed in a more general education program such as GED preparation. One of the facilities in the sample reported that it had no Title I funds to expend. A follow-up contact revealed that this facility now paid for the Title I N or D program with Title I Handicapped funds. Incidentally, this institution reported spending nearly $11,600.00 per student for education. This was the largest amount spent per pupil in any institution, exceeding the sum of the Title I and non-Title I per-pupil expenditures at any other insitution in the sample.

One of the facilities reported spending no Title I funds on reading instruction. Another reported that it did not provide a non-Title I-funded reading program. Three institutions did not fund mathematics instruction with Title I funds, while two others did not fund mathematics instruction with non-Title I funds.

Distribution of Teaching Assignments

Only teachers of the students selected for testing were asked to complete questionnaires, and thus the teachers who completed questionnaires do not necessarily represent the population of teachers at institutions.

Three hundred and eleven teachers at the 30 randomly selected institutions returned questionnaires. Table IV-9 shows the distribution of teaching assignments across subject matters. Because some teachers had multiple teaching assignments, the number of teachers in Table IV-9 exceeds 311. Nearly 70 percent of the classes taught were in regular education. Of all classes, 29 percent were in reading or language arts, 26 percent were in mathematics, 20 percent were in other academic subject matters (e.g., science, social studies) and 25 percent were in non-academic areas (e.g., vocational skills).

Table IV-9. Distribution of Teaching Assignments
(30 Randomly Selected Institutions)

Subject Matter and Program	N*
Reading/Language Arts	
Title I	68
Regular Education	90
Mathematics	
Title I	52
Regular Education	86
Other Academic	
Title I	22
Regular Education	84
Non-Academic	
Title I	28
Regular Education	104

*Some teachers had multiple assignments. The total number of teachers was 311.

Table IV-10 shows the distribution of teaching assignments in basic skills. In this table each teacher is counted once only. Of the 189 teachers who teach reading and/or mathematics classes, 34 percent teach only reading (or language arts), 24 percent teach only mathematics, and 42 percent teach both reading and mathematics classes. Over a third of these teachers teach only Title I classes, half teach only regular education classes, and 8 percent teach both Title I and regular education classes.

Table IV-10. Distribution of Basic Skills Assignments
(30 Randomly Selected Institutions)

Subject Matter and Program	N*
Reading Only	
Title I Only	30
Regular Education Only	31
Title I and Regular Education	3
Total	64
Mathematics Only	
Title I Only	14
Regular Education Only	28
Title I and Regular Education	3
Total	45
Reading and Mathematics	
Title I Only	25
Regular Education Only	46
Title I and Regular Education	9
Total	80

*Each teacher is counted once only.

Teaching Practice and Classroom Support

Only responses by the 189 teachers of basic skills (reading or language arts and/or mathematics) were included in the following analyses.

Because many of the variables on the teacher questionnaires were correlated, preliminary analyses were performed to obtain a parsimonious set of variables that could then be related to student outcomes. Many analyses of student outcomes were to be performed using data aggregated to the institution level, so analyses of teacher variables were selected that would differentiate among institutions. Stepwise discriminant analysis was selected as most appropriate for this analysis. This method of analysis finds the smallest set of variables that distinguishes among institutions. Because the full set of variables contained redundant information, and some of the variables were not useful in discriminating among institutions, the stepwise discriminant analyses reduced the total set of variables to a manageable number. In Appendix E the institutional means and standard deviations are given for the full set of variables on the teacher questionnaires that showed some variation across institutions.

The reduced set of variables resulting from the stepwise discriminant analysis is described in Table IV-11. This table presents means and standard deviations for each of the variables in the reduced set. In this table, teacher responses were aggregated to the institution level; these institutional aggregates were used to determine means and variances. Thus, institutions were weighted equally. The variables resulting from the discriminant analysis of responses by Title I and regular education teachers are listed in the first section of Table IV-11. The variables resulting from the analysis of questionnaires completed by Title I only teachers are listed next. This table is based on the mean response for each institution. Thus, the results reflect the variation across institutions, but not within any institution. Within-institution variation is considered later in this section. The variables in the following discussion proved to discriminate among institutions.

As may be seen in Table IV-11, teachers rated cooperation between students and custody staff as good, on the average. Teachers reported that facility administrators were usually supportive of their work. The third variable in this set was whether a teacher possessed a state teaching credential or license. Most teachers reported that they did have such a credential. Among the responses by Title I teachers about their Title I reading or mathematics classes, teachers reported that they received three to four hours of assistance per week from special reading or math teachers. They were assisted by a counselor or psychologist several times a year on the average. Title I reading students used instructional materials such as textbooks and workbooks weekly; Title I mathematics students used such materials only once or twice a month. Title I mathematics students used special equipment such as feedback teaching machines every few months on the average. Title I reading and mathematics teachers tended to use lesson plans that provided broad objectives and categories of activities for each day's instruction rather than detailed lesson plans or no lesson plans at all. Title I teachers reported that they gave students feedback in the form of a letter or number grade about once a month, on the average. Title I reading teachers received assistance from a media specialist about once a year. Title I mathematics teachers received about 13 hours of assistance per week from a teacher aide or adult assistant.

Although these figures show how teachers responded on the average, there was substantial variation among institutions on these variables. Appendix F contains the means and standard deviations of each of the variables in Table IV-11 for each institution. Some of the variables show wide variation across all institutions; others show uniformity across most institutions, with a few institutions standing apart from the rest.

Table IV-12 shows the range of teacher responses within institutions. Although some institutions show a wide range, as indicated by the 'highest standard deviation', most institutions show considerable agreement among teachers, as indicated by the small mean standard deviations.

Table IV-11. Distribution of Mean Teacher Responses on Variables Found to Distinguish among Institutions (Data are average responses of teachers at the 30 randomly sampled institutions.)

Teacher Response	n	Mean	s.d.	Range of Variables	
				Low	High
All Reading and Mathematics Teachers					
Cooperation between students and custody staff.	28	2.8	0.5	1	4
Teacher's perception of support from facility administrators.	29	3.3	0.5	1	4
Possession of state teaching credential or license.	29	1.1	0.2	1(yes)	2(no)
Title I Reading Teachers					
Amount of support from special reading teacher (hours per week).	28	3.1	7.7	Continuous	
Amount of assistance from counselor, psychologist	27	2.3	1.2	1	4
Amount of student use of instructional materials.	26	33.0	5.7	7	42
Amount of advance planning.	28	2.0	0.8	1	3
Frequency of grading students.	28	3.5	1.9	1	6
Amount of assistance from media specialist.	27	1.7	1.1	1	4
Title I Mathematics Teachers					
Amount of assistance from special mathematics teacher (hours per week).	17	4.1	10.0	Continuous	
Amount of assistance from counselor, psychologist.	23	1.9	1.1	1	4
Amount of student use of instructional materials.	22	21.7	3.8	5	30
Amount of advance planning	24	1.9	0.8	1	3
Frequency of grading students.	23	4.1	2.0	1	6
Amount of support from teacher aide or assistant (hours per week).	23	13.3	15.2	Continuous	
Amount of student use of special instructional equipment	24	13.4	7.1	5	30

Table IV-12. Average Variation of
Teacher Responses Within Institutions

Teacher Response	Range in S.D.		
	Lowest	Highest	Mean S.D.
All Reading and Mathematics Teachers			
Cooperation between students and custody staff.	.0	1.2	.6
Teacher's perception of support from facility administrators.	.0	1.4	.6
Possession of state teaching credential or license.	.0	.7	.2
Title I Reading Teachers			
Amount of support from special reading teacher (hours per week)	.0	21.2	2.1
Amount of assistance from counselor, psychologist.	.0	1.4	.2
Amount of student use of instructional materials.	.0	12.0	1.8
Amount of advance planning.	.0	1.4	.3
Frequency of grading students.	.0	3.5	.7
Amount of assistance from media specialist.	.0	2.1	.2
Title I Mathematics Teachers			
Amount of assistance from special mathematics teacher (hours per week).	.0	17.3	.6
Amount of assistance from counselor, psychologist.	.0	1.0	.1
Amount of student use of instructional materials.	.0	5.7	.6
Amount of advance planning.	.0	1.0	.1
Frequency of grading students.	.0	3.5	.4
Amount of support from teacher aide or assistant (hours per week).	.0	20.5	1.9
Amount of student use of special instructional equipment.	.0	6.8	.8

Comparison of Selected Teaching Practices With Previous Findings

Information on several variables was obtained in a previous study of compensatory education, the NIE Compensatory Education Study. The NIE study examined compensatory education in public schools for students in kindergarten through grade eight. Table IV-13 compares how teachers planned lesson content and assigned instructional tasks in the NIE study to the reported behavior of teachers in the present study. In the NIE study, most teachers planned what to teach on the basis of standardized achievement test scores and almost half planned on the basis of diagnostic test scores. Relatively few teachers in that study used criterion or objective referenced test scores or the student's age to plan course content. In the present study most teachers used standardized achievement and diagnostic test scores to plan what to teach. Almost half used criterion or objective referenced test scores. Again, relatively few used the student's age, although more emphasis was placed on age in the present study than in the NIE study.

That relatively few teachers based instructional content on the student's age shows that, in general, planning by compensatory education teachers was sensitive to the instructional needs of students in compensatory education, who usually perform below the level of their peers in regular education programs. One possible explanation for the greater use of students' age in the present study is that teachers may have interpreted "use of student age" to mean use of materials suitable for older students. As the average age of the Title I students was far above their achievement level, a difficult task for Title I reading teachers was to find reading materials that were aimed at low-achieving adult readers. The greater use of diagnostic and criterion or objective referenced test scores in the present study may be due to the greater availability to institutional teachers of test results. At most of the institutions, entering residents are given a battery of tests at a diagnostic center before being assigned to an educational program. These tests are often diagnostic or objective referenced.

Table IV-13. Comparison of Selected Responses of Reading Teachers to Previous Findings

Variable	Present Findings	Previous Findings*
Percent of teachers who plan course content based on		
Standardized achievement test scores	70	77
Diagnostic test scores	79	46
Criterion or objective referenced test scores	42	19
Student's age	24	15
Percent of time tasks are assigned to**		
Whole class	24	17
Small groups of students	14	32
Individual students	52	38

Table IV-13 shows that teachers of compensatory education of confined youth emphasized individualized instruction more, and small group instruction less, than was found in the NIE study. In confined settings the range in student ages and previous achievement is probably greater than the range in public school compensatory education settings. To respond to student needs, then, teachers of confined youth may view assignment of tasks to individual students as the best way to match instruction to the student's achievement level.

*Evaluating Compensatory Education: An Interim Report on the NIE Compensatory Education Study (Washington, December 1976).

**That the figures do not add up to 100% may be due to teachers' including preparation time and breaks as part of instructional time, but not included in the categories listed.

Assignment of Students to Educational Programs

Federal guidelines direct that only students eligible for Title I services (those under 21 years of age, lacking a high school diploma or GED) may be enrolled in a Title I program. It is further specified that students most in need of basic skills instruction should have the highest priority for receipt of Title I services. However, the guidelines on how to determine which students are most in need of Title I instruction are not specific, and institutions vary on how students are assigned to Title I instruction or regular education alone. This section describes how students are selected for participation in educational programs and how they are assigned to particular programs.

The educational program administrator at each institution was asked to describe how students were selected for participation in an educational program. Table IV-14 summarizes their responses. Nearly all institutions for neglected or

Table IV-14. Basis for Selecting Students to Participate in Educational Programs

Basis for Selection	Number of Institutions
Willingness to participate	4
Diagnostic test scores	3
Diagnostic test scores and institutional staff recommendations	2
Willingness to participate and diagnostic test scores	1
Willingness to participate and institutional staff recommendations	3
Willingness to participate, institutional staff recommendations and diagnostic test scores	4
All residents participate	23
Total	40

delinquent youth assigned all students to an educational program. Most of the adult institutions had a mechanism for selecting students, such as varying combinations of willingness to participate, recommendations by institutional staff, and diagnostic test scores.

Both the educational program administrator and Title I reading and mathematics teachers were asked to describe how students were assigned to Title I instruction. Their responses are given in Table IV-15. According to the educational program administrators, nearly half of the institutions assigned all students or all eligible students to Title I programs. The remaining institutions used a mixture of test scores, staff recommendations, and student willingness to

Table IV-15. Basis for Assigning Students to Title I Programs

Basis for Assignment as Reported by	Number of Institutions
Educational Program Administrators	
Test scores only	6
Staff recommendations	1
Test scores and staff recommendations	7
Willingness to participate	4
Willingness to participate and test scores	1
Willingness to participate, staff recommendations and test scores	3
All eligible students participate	13
All students participate	5
Total	40
Title I Teachers	
Test scores only	6
Test scores and teacher or administrator judgment	20
Teacher or administrator judgment and student request	1
Test scores, teacher or administrator judgment, and student request	11
Total	38*

*Title I teachers at two institutions did not respond to this question.

participate in order to assign students to Title I instruction. The reports of Title I teachers generally agreed with those of the educational program administrator.

It is difficult to determine how systematically students were assigned to Title I instruction in institutions where some selection procedure was operative. There are several features operating in these institutional environments that may work against systematic assignment. First, that student request for participation (usually related to the institution's system of rewards) was sometimes the basis for assignment suggests that assignment may not always have been based on need. Especially among adult institutions, where residents by law (and at least theoretically if not pragmatically) can either elect or refuse program participation, educational program participation is likely to be based on individual motivation rather than on need. Because most students were over 16 years of age and were therefore not required to enroll in an educational program, the institution could not mandate participation in a specific educational program. Students greatly in need of Title I instruction who wanted to enroll in a regular education program, but who did not want to receive Title I instruction, could not be forced to do so. Conversely, students who had less need for Title I instruction but who desired to participate could not be denied instruction if space were available.

To better assess whether students were systematically assigned to educational programs, institutions using test scores as the basis for assignment were asked to specify the value of the cutoff score, if one was used. Of the six institutions that used test scores as the only basis for assignment, four institutions specified a cutoff score. This cutoff score was usually a certain number of years below grade level. These four institutions were then asked to provide their own test scores for the students tested in the present study. As the program designation of each student had already been provided, examination of

the test scores and the cutoff score showed whether the institution systematically used their test scores to assign students to educational programs.

In three of the four institutions the score distributions of Title I and regular education students overlapped--many students assigned to Title I instruction had scored higher than students assigned to regular instruction. There was little evidence that students had been assigned according to the cutoff specified. At the remaining institution, students in Title I programs did score lower than students receiving only regular education instruction, with a few exceptions. Of the four institutions which maintained that students were assigned to educational programs according to a cutoff score, only one showed evidence of assigning the most needy students--as measured by test scores--to the Title I program. Thus, out of 40 institutions, only one appeared to assign students to Title I using test scores in a clearly systematic manner, i.e., in only one institution was it clear that the 'neediest of the needy' students, as measured by test scores, were being served.

Although only one institution used test scores systematically, other institutions may have used a combination of factors, including teacher judgment, to systematically assign students to educational programs.* The data collected here do not address this possibility. There are published opinions that suggest that one should not rely entirely on test scores to classify students. Tallmadge (1978) argues that a combination of classroom records, teacher judgment, and achievement test scores provides a more valid and reliable standard than single measures. Whether institutions use single or combination measures to assign students to programs systematically has yet to be determined.

Systematic assignment procedures are important because they provide clear guidelines for defining the population of students to be served by compensatory education programs, and because they enable one to make an accurate evaluation of

*There is some evidence that in certain settings teacher judgments may agree with test scores. Kenoyer (1978) found that teacher judgment of student need for compensatory education (yes or no decision) services agrees moderately with scores on the CTBS.

the influence of such programs on achievement. This procedure is described under the name "Systematic Allocation Model" in the Handbook for Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth (Webb and Keesling, 1978). The fact that almost one-half of the institutions provide Title I services to all students, however, may make it unlikely that this model will be widely adopted. While it might be argued that institutions could implement better assignment to Title I treatment procedures that would allow for a more accurate evaluation of the impact of Title I on achievement gains, it could also be argued that the entire population of students confined in state institutions is in great need of basic skills instruction.

Provision of Instruction in Basic Skills

The teachers who offered basic skills instruction to the sampled students at each institution were requested to fill out participation records for each sampled student in each class. One such record was to be filled out prior to each of the four testing periods. The record was to indicate the number of minutes the class was scheduled to be held, the number of minutes actually held, and the number of minutes each student attended. The form to be filled out requested additional data on the amount of time each student spent working in groups of various sizes (or alone) and the amount of time during which the professional staff, aides, peers, or teaching machines were the source of the instructional direction given the student. The logistics of distributing the forms and the complexity of the forms themselves led to the omission of much information at several institutions. For example, because SDC could not obtain student names, the data for a student were to be labeled with the identification number that was assigned from our roster lists. Many teachers were unable to locate these numbers because the rosters were not conveniently available to them, or because they misunderstood our request and put in an institutional identification number.

The first form was markedly affected by these errors since the procedure was new. But even the fourth forms were not filled out well because tracking students became more complex as there were four rosters of names to be included at that point, because some pupils had left the institution and, because many institutions added large numbers of students to the last roster.

From the forms returned at the second and third data collection points, usable data was recovered from 17 of the 30 randomly selected institutions. These data are summarized in Table IV-16. The data are for students who were selected for the study on the first or second visit to institutions and who were listed on at least one participation record from the second visit and one from the third visit. Each participation record refers to the week prior to the week of the respective data collection visit. While it may be too optimistic to treat these data summaries as reasonable estimators of population means, they reveal apparent characteristics of Title I services in terms of possibly limited exposure which clearly merit further study. The extent to which interruptions to the provision of instructional services are due to institutional constraints as opposed to the lack of student motivation or willingness to participate cannot be readily assessed based on data obtained in Phase II. SDC will be conducting such a study on a small scale in the next year. In that study the data on time-of-instruction over 12 weeks will be obtained by trained observers, with a much less complicated supplementary record to be kept by the teaching staff.

One of the striking features of the data in Table IV-16 is the propensity of students in Title I programs in both reading and mathematics to fail to attend the scheduled Title I classes. Attendance at all reading classes (Title I or regular education) is between 90 percent and 100 percent for all groups of students, except that those in Title I supplementary programs in both reading and mathematics attend only 70 percent of the offered Title I reading class minutes. This measure is undoubtedly a combination of student motivation and willingness to participate in educational programs and institutional scheduling. Additionally, the discrepancy between the number of scheduled minutes and the number of minutes class was actually held is also large for this group of students, and provides a clearer indicator of the extent to which institutional constraints or other institutional activities interrupt instruction (see Table IV-17). Institutions running Title I supplementary programs in both basic skills areas (the majority of the institutions sampled do so) tend to pull students away from the Title I reading class, or even to cancel a class, rather than disturb other institutional activities.

Table IV-16. Minutes of Instruction Per Week in Basic Skills Subjects (Weighted means from 17 randomly sampled institutions)

Minutes of Instruction Per week in	Student is Enrolled in a Program of		
	Regular Education Instruction Only	Regular Education Plus Title I Supplement in Reading and Mathematics	Regular Education Plus Title I Supplement in Reading only
	(n=58)	(n=197)	(n=68)
Non-Title I Reading			
Scheduled	250	90	65
Held	230	85	60
Attended	230	80	60
Title I Reading			
Scheduled	50	165	190
Held	50	135	185
Attended	50	95	165
Non-Title I Mathematics			
Scheduled	160	190	195
Held	150	190	185
Attended	115	70	155
Title I Mathematics			
Scheduled	*	230	*
Held	*	225	*
Attended	*	75	*
Total Reading			
Scheduled	300	255	255
Held	280	220	245
Attended	280	175	225
Total Mathematics			
Scheduled	160	420	195
Held	150	415	185
Attended	115	145	155

* Fewer than 10 minutes per week were reported, on average.
Numbers are rounded to nearest 5 minutes.

Table IV-17. Percentages of Scheduled Instructional Times That Were Held and Attended

Instruction in	Student is Enrolled in a Program of		
	Regular Education Instruction Only	Regular Education Plus Title I Supplement in Reading and Mathematics	Regular Education Plus Title I Supplement in Reading Only
Reading			
Percent of scheduled instructional time which was actually held	93	86	96
Percent of scheduled instructional time which was actually attended	93	69	88
Mathematics			
Percent of scheduled instructional time which was actually held	94	99	95
Percent of scheduled instructional time which was actually attended	72	35	79

While the Title I mathematics classes do not show the same effect with respect to scheduled vs held minutes, the individual student attendance is clearly very poor--being less than 1/2 the rate of attendance by other groups in regular education mathematics classes. The students receiving Title I supplementary services in mathematics and reading do not attend the regular mathematics classes at a greater rate than they attend the Title I supplementary classes. As a consequence, although they are offered nearly twice the instruction in mathematics as their peers who are in Title I supplementary programs in reading only, they receive less actual exposure!

Another interesting feature of this table is the fact that the students in programs of regular education instruction only are offered nearly as much reading instruction through their non-Title I reading classes as the Title I students are offered in combined non-Title I and Title I reading classes. Because the regular education only students are present for more of the instruction, they receive more instruction from non-Title I reading classes than Title I students receive from the combination of non-Title I and Title I classes. The fact that non-Title I students receive Title I reading instruction gives them nearly an additional hour of actual instruction in reading each week than their Title I peers receive.*

While these data are suggestive, it is important to note that many programs of instruction at institutions are very individualized. Thus, while the time in a classroom was attributed to Title I or regular education, based upon the teacher's designation of whether the class was a Title I class or otherwise, an individual student could have been pursuing non-Title I instruction in a Title I class, or vice-versa. Bearing this caution in mind (as well as the small numbers of cases upon which the tabled data are based), it seems that students in programs which attempt to remediate both reading and mathematics deficiencies do not receive the full amount of the supplementary services

*Note that this is an average allocation of time. Some non-Title I students are in Title I classes very often, others not at all.

offered. They seem to be unable (or unwilling?) to attend any mathematics classes and their attendance at Title I reading classes is low. The loss of 30 minutes per week of supplementary instruction in reading due to not holding scheduled classes suggests that, to some extent, the low rate of receipt of services in this group is due to institutional effects perhaps beyond the control of the students.

These data on instructional time do not compare well to the data on costs (Table IV-8). Especially significant is the fact that Title I supplementary services in mathematics cost twice the amount of the regular education mathematics program on a per-pupil basis, yet the intended recipients of these services attend only one-third of the instructional minutes in mathematics purchased by the Title I program. Unfortunately, the data collected suggest neither the cause nor the remedy for this mismatch between offered services and their utilization.

In the next chapter data are presented on the achievement levels and attitudes of the sampled students. In addition, the relationships of institutional and program characteristics to achievement and attitude variables are explored.

CHAPTER V. THE RELATIONSHIP OF THE CHARACTERISTICS OF INSTITUTIONS AND EDUCATIONAL PROGRAMS TO THE ACHIEVEMENT AND ATTITUDES OF INSTITUTIONALIZED YOUTH

PRECIS OF CHAPTER V

This chapter presents data on the achievement and attitude measures obtained at the four waves of testing. Standardized achievement test results are used to indicate that students in institutions for neglected or delinquent youth are performing at a level very much behind that expected of 16 to 18 year olds. Indeed, their performance is comparable to that of 4th through 6th graders. Some additional data on this low level of achievement came from a comparison of certain parts of the mathematics domain referenced tests to certain parts of a test given to 11th graders in the state of Florida. These comparisons, reported in Appendix I, show institutionalized neglected or delinquent youth to be very much behind their age peers in the performance of basic arithmetic skills.

Results for the domain referenced tests are presented in the next section of this chapter. They show that, overall, there is very little change in the achievement levels of Title I or non-Title I students during the twelve weeks between the first and third testings. The average performance dropped considerably from the third to the fourth testing. We have chosen to interpret this as an effect of repetitious testing rather than a true reduction of skill level. At the first testing, the scores of Title I students were below those of non-Title I students. Generally speaking, these two groups of students progressed at the same rate during the course of the study. Some institutions did seem to have more success than others, however.

The attitude measures showed similar results. Changes were small and did not seem to occur at different rates for Title I and non-Title I students.

Because the Title I group is generally younger and has lower levels of achievement and less positive attitudes, it is difficult to establish reasonable expectations for the Title I student group and to know how they would fare if Title I programs were not available. The presence of Title I funds may be enabling

institutions to provide services to a group which would not be served by or would not be able to benefit from the services provided by regular, state-funded education programs.

Gains in achievement did not appear to be related to measured student attitudes. To gain some insight into other factors which might affect the attitude and achievement outcomes measured in this study, such as the characteristics of institutions and educational programs, each of the attitude and achievement measures taken at the third testing was regressed on the corresponding measure taken at the first testing, and the residuals were correlated with each of the measures of institutional and program characteristics. The most promising of the measures of characteristics identified in this fashion were assessed further by performing a multiple regression analysis. A set of models resulted which is relatively parsimonious and accounts reasonably well for the observed data. Treating the coefficients in the models more as indicators of direction in the data--rather than as firm estimates of parameters in theoretical models--it seems reasonable to conclude that institutions which are able to devote time and effort to the achievement and attitude needs of their students should be more successful than others in effecting changes in these student characteristics. Specifically, smaller programs, with low rates of staff turnover, in institutions which consider custody of secondary importance (behind education or counseling) should be more successful.

Finally, the analysis of the ten purposively selected sites (where selection was based on Phase I data collectors' observations of program characteristics such as strong staff commitments to instruction, strong administrative leadership, student involvement in learning, positive staff-student interactions, etc.) revealed no differences between the purposively selected or randomly selected sites. While no separate reporting of the characteristics of the purposively selected sites is documented herein, these sites were represented in the exploratory analyses that investigated the relationship of institutional and educational program characteristics to achievement.

CHAPTER V

This chapter presents the levels of achievement and attitudes as measured by the instruments presented in Chapter III. In addition, the program characteristics presented in Chapter IV are related to certain achievement and attitude measures.

STANDARDIZED ACHIEVEMENT TEST RESULTS

In this section we present the results of the testing conducted with standardized achievement tests. These results enable us to compare the performance of students in state institutions to the performance of students (of very different ages but similar levels of skill) who are attending regular schools. In general, these comparisons show the entire population of neglected or delinquent students to be performing well below grade level.

Practical Achievement Test

Table V-1 shows the results of this test of functional literacy, administered at the second visit to sites, and the results for comparison groups of non-institutionalized youth obtained from the Sustaining Effects Study being conducted by System Development Corporation. The scores used were the raw numbers of correct responses out of a total of 30 items. The internal consistency reliability of this test among the neglected or delinquent students was .93, using coefficient alpha. Recall that Title I neglected or delinquent students are about 16 1/2 years old and that non-Title I neglected or delinquent students are about 18 1/4 years old, on the average. The neglected or delinquent Title I students perform less well than non-institutionalized 5th-grade students who do not receive compensatory education. The neglected or delinquent non-Title I students perform on a par with the non-institutionalized 6th graders who are not receiving compensatory education.

Comprehensive Tests of Basic Skills

Level 2 of the Comprehensive Tests of Basic Skills was given on the fourth visit to sites. The coefficient alpha internal consistency reliability was .97 for both the mathematics form (73 items) and for the reading form (85 items). Table V-2 gives the results for the reading test and Table V-3 gives the results for the mathematics test. Each table shows results for comparison groups

Table V-1. Results for Practical Achievement Scale
(Scores are Raw Numbers of Correct Items)

Practical Achievement Scale Results in 30 Randomly Selected Neglected or Delinquent Institutions		
Title I Students	Mean	16.74
	s	8.29
	n	635
Non-Title I Students	Mean	21.51
	s	7.65
	n	333

Summary of Sustaining Effects Study			
Students in Schools With No Title I Funding			
		Grade 5	Grade 6
Students served by other (non-Title I) Compensatory Education program	Mean	13.95	15.80
	s	6.14	5.99
	n	538	535
Students not served by any Compensatory Education program in schools offering non-Title I Compensatory Education services	Mean	20.10	22.43
	s	6.14	5.51
	n	2628	4321
Students not served by any Compensatory Education program in schools having no Compensatory Education services	Mean	19.30	21.25
	s	6.67	6.09
	n	1242	1656
Students in Schools With Only Title I Compensatory Education			
Title I Students	Mean	14.62	16.01
	s	5.51	6.38
	n	418	340
Students not served by any Compensatory Education program	Mean	20.07	21.84
	s	5.94	5.67
	n	2195	2243

Table V-2. Results for Comprehensive Test of Basic Skills
(Level 2 Reading Test Grades 4 and 5)
Scores are Percentiles of the National Norm Distribution

Results From 30 Random Neglected or Delinquent Institutions

		Using 4th-Grade Norms	Using 5th-Grade Norms
Title I Students	Mean	48.68	38.09
	s	34.09	31.63
	n	737	737
Non-Title I Students	Mean	63.77	53.31
	s	33.48	33.73
	n	371	371

Summary of Sustaining Effects Study

Students in Schools With No Title I Funding

		Grade 4	Grade 5
Students served by other (non-Title I) Compensatory Education program	Mean	32.55	27.04
	s	23.86	20.61
	n	527	542
Students not served by any Compensatory Education program in schools offering non-Title I Compensatory Education services	Mean	60.92	60.05
	s	27.50	26.68
	n	2487	2640
Students not served by any Compensatory Education program in schools having no Compensatory Education services	Mean	52.50	[illegible]36
	s	28.94	29.62
	n	1184	1250

Students in Schools With Only Title I Compensatory Education program

Title I Students	Mean	28.85	30.60
	s	20.25	19.94
	n	500	418
Students not served by any Compensatory Education program	Mean	56.76	58.63
	s	26.56	26.07
	n	2126	2204

Table V-3. Results for Comprehensive Test of Basic Skills
(Level 2 Mathematics Test Grades 4 and 5)
Scores are Percentiles of the National Norm Distribution

Results in 30 Randomly Selected Neglected or Delinquent Institutions

		Using 4th-Grade Norms	Using 5th-Grade Norms
Title I Students	Mean	47.34	34.86
	s	35.11	31.91
	n	718	718
Non-Title I Students	Mean	63.26	51.05
	s	35.00	34.86
	n	364	364

Summary of Sustaining Effects Study

Students in Schools With No Title I Funding

		Grade 4	Grade 5
Students served by other (non-Title I) Compensatory Education program	Mean	35.65	28.98
	s	26.49	24.71
	n	362	404
Students not served by any Compensatory Education program in schools offering non-Title I Compensatory Education services	Mean	56.28	56.79
	s	29.29	28.35
	n	2646	2773
Students not served by any Compensatory Education program in schools having no Compensatory Education services	Mean	50.22	52.11
	s	28.02	29.71
	n	1183	1247

Students in Schools With Only Title I Compensatory Education program

Title I Students	Mean	35.60	31.32
	s	24.70	21.80
	n	220	246
Students not served by any Compensatory Education program	Mean	52.85	54.74
	s	27.45	27.14
	n	2439	2416

110

sampled in the Sustaining Effects Study. All raw scores were converted to percentiles using the publisher's norm tables for grades four and five.

Generally speaking, the neglected or delinquent Title I students perform similarly to 4th-grade non-institutionalized students who do not receive compensatory education. Neglected or delinquent non-Title I students perform on a par with non-institutionalized 5th-grade students not receiving compensatory education. The very large age difference between the institutionalized students and the students sampled in the Sustaining Effects Study suggests that neglected or delinquent students are very far behind their non-institutionalized age peers.

The gap separating the neglected or delinquent Title I students from the neglected or delinquent non-Title I students is about one-half of the standard deviation within either group. This suggests that about 30 percent of the non-Title I students perform less well than the average Title I student in state institutions. Generally speaking, the neediest students are selected for Title I services, however, a large proportion of the students who are not served are needier than the average pupil who is served.

RESULTS OF DOMAIN REFERENCED TESTS

A random sample of approximately 60 students was selected at each institution in the sample at the time of the first testing. Some of these students were present for testing throughout all four waves of testing; others were present for fewer than four administrations. At each successive wave of testing new students were added to the roster with some being selected for testing. Though some of these added students were present and were tested through the fourth test administration, many were not. A group of students who were selected into the sample at the same time is called a "cohort". Thus, "Cohort 1" is the group of students that was selected and tested on the first visit to the institutions. In the subsequent tables, the performance of each cohort is described separately. Each of these cohorts is further subdivided into groups according to the sequence of testings for which data were obtained from the group. Thus, Cohort 1 is divided into four groups: those who took only the first test, those who took tests at the first and second administrations, those who took tests at the first

three administrations, and those who took all four tests. The other cohorts have fewer groups because there are fewer testing patterns for those cohorts. Each group in each cohort is further divided into those students who received Title I instruction and those who received only regular instruction.

The description of performance in Tables V-4 and V-5 shows several results. The most striking result is that the performance of nearly every cohort of students declined markedly at the fourth wave of testing. Regardless of when students entered the test cycle, they showed a poorer performance at the fourth wave of testing than at any other wave. This finding is equally strong for Title I students and students who received only regular instruction. It is unreasonable to suggest that students forgot what they mastered six weeks previously. Rather, this decline seems to be the result of a negative reaction to repeated six-week testings. Put simply, students were tired of taking tests. Test administrators indicated that students were aware that the fourth wave of testing marked the end of the testing phase, and were more obstreperous than at earlier test administrations. This negative reaction extended even to newcoming students at the fourth wave, who performed at the level of the other students at the fourth test administration.

The decline in scores at the fourth wave seems to be related to increasing numbers of omitted items, as shown in Table V-6. The increased number of omissions does not, however, completely account for the decline in scores. Another part of the decline may have been due to increasing numbers of random responses, but this aspect of the score decline was not examined, since it would have required a clinical judgment of each test protocol.

Although the increasing rate of omissions does not disprove the hypothesis that score declines are due to forgetting previously acquired skills, declining motivation seems to be a more plausible explanation of the data. It should be noted that standard deviations generally increased at the fourth wave of testing. This finding supports the notion that the task is not being taken seriously by a substantial number of students at the later testing. Because the test scores from the fourth wave may be symptomatic of fatigue and boredom, rather than the forgetting of previous learning, further analyses of achievement will concentrate on the first three waves of test data.

Table V-4a. Proportion Correct on Arithmetic DRT for Students First Tested at Wave 1 in Randomly Selected Institutions

Students Tested at:	Results for:	Title I Plus Regular Education				Regular Education Only			
		Wave 1	Wave 2	Wave 3	Wave 4	Wave 1	Wave 2	Wave 3	Wave 4
Wave 1 Only	Mean =	.63	--	--	--	.68	--	--	--
	s =	.17	--	--	--	.22	--	--	--
	n =	223	--	--	--	94	--	--	--
Waves 1 and 2	Mean =	.68	.67	--	--	.71	.71	--	--
	s =	.17	.20	--	--	.16	.18	--	--
	n =	111	111	--	--	70	70	--	--
Waves 1, 2, and 3	Mean =	.60	.61	.60	--	.74	.75	.74	--
	s =	.20	.21	.21	--	.16	.18	.17	--
	n =	88	88	88	--	38	38	38	--
All four waves	Mean =	.65	.64	.65	.56	.70	.72	.72	.66
	s =	.17	.17	.17	.20	.16	.15	.19	.20
	n =	140	140	140	140	70	70	70	70

The means and standard deviations in this table are estimated by weighting cases for probability of selection.

Table V-4b. Proportion Correct on Arithmetic DRT for Students First Tested After Wave 1 in Randomly Selected Institutions

Results for:		Title I Plus Regular Education			Regular Education Only		
		Wave 2	Wave 3	Wave 4	Wave 2	Wave 3	Wave 4
Students First Tested at Wave 2							
Students Tested at:							
Wave 2 Only	Mean =	.64	--	--	.80	--	--
	s =	.16	--	--	.17	--	--
	n =	55	--	--	34	--	--
Waves 2 and 3	Mean =	.59	.60	--	.79	.80	--
	s =	.16	.17	--	.14	.12	--
	n =	26	26	--	23	23	--
Waves 2, 3, and 4	Mean =	.64	.62	.54	.80	.79	.71
	s =	.13	.14	.18	.20	.19	.22
	n =	60	60	60	25	25	25
Students First Tested at Wave 3							
Students Tested at:							
Wave 3 Only	Mean =	--	.68	--	--	.74	--
	s =	--	.16	--	--	.12	--
	n =	--	105	--	--	62	--
Waves 3 and 4	Mean =	--	.65	.60	--	.69	.61
	s =	--	.19	.20	--	.17	.21
	n =	--	114	114	--	76	76
Students First Tested At Wave 4							
Students Tested Only at Wave 4	Mean =	--	--	.58	--	--	.64
	s =	--	--	.20	--	--	.18
	n =	--	--	145	--	--	103

The means and standard deviations in this table are estimated by weighting cases for probability of selection.

Table V-5a. Proportion Correct on Literal Comprehension DRT for Students First Tested at Wave 1 in Randomly Selected Institutions

Students Tested at:	Results for:	Title I Plus Regular Education				Regular Education Only			
		Wave 1	Wave 2	Wave 3	Wave 4	Wave 1	Wave 2	Wave 3	Wave 4
Wave 1 Only	Mean =	.63	--	--	--	.76	--	--	--
	s =	.25	--	--	--	.19	--	--	--
	n =	249	--	--	--	95	--	--	--
Waves 1 and 2	Mean =	.67	.63	--	--	.73	.72	--	--
	s =	.27	.27	--	--	.21	.24	--	--
	n =	112	112	--	--	71	71	--	--
Waves 1, 2, and 3	Mean =	.57	.55	.54	--	.69	.65	.73	--
	s =	.28	.28	.29	--	.25	.24	.24	--
	n =	96	96	96	--	41	41	41	--
All Four Waves	Mean =	.67	.62	.64	.57	.76	.71	.76	.71
	s =	.22	.26	.26	.29	.24	.28	.21	.25
	n =	133	133	133	133	62	62	62	62

The means and standard deviations in this table are estimated by weighting cases for probability of selection.

Table V-5b. Proportion Correct on Literal Comprehension DRT for Students Tested After Wave 1 at Randomly Selected Institutions

Results for:		Title I Plus Regular Education			Regular Education Only		
		Wave 2	Wave 3	Wave 4	Wave 2	Wave 3	Wave 4
Students First Tested at Wave 2							
Students Tested at:							
Wave 2 Only	Mean =	.60	--	--	.80	--	--
	s =	.28	--	--	.18	--	--
	n =	55	--	--	31	--	--
Waves 2 and 3	Mean =	.57	.54	--	.85	.86	--
	s =	.22	.24	--	.13	.15	--
	n =	32	32	--	27	27	--
Waves 2, 3, and 4	Mean =	.53	.51	.45	.69	.69	.67
	s =	.26	.27	.27	.26	.26	.29
	n =	69	69	69	33	33	33
Students First Tested at Wave 3							
Students Tested at:							
Wave 3 Only	Mean =	--	.65	--	--	.67	--
	s =	--	.30	--	--	.30	--
	n =	--	119	--	--	75	--
Waves 3 and 4	Mean =	--	.65	.62	--	.71	.69
	s =	--	.26	.26	--	.25	.27
	n =	--	125	125	--	75	75
Students First Tested at Wave 4							
Students Tested Only at Wave 4	Mean =	--	--	.58	--	--	.67
	s =	--	--	.28	--	--	.25
	n =	--	--	149	--	--	110

The means and standard deviations in this table are estimated by weighting cases for probability of selection.

Table V-6. Mean Proportion Correct and Mean Proportion Omitted on Achievement Tests for Students Whose Performance Decreased From Wave 1 to Wave 4 (Students from randomly selected institutions only)

Test Program	n	Wave 1		Wave 2		Wave 3		Wave 4	
		Prop Correct	Prop Omicted	Prop Correct	Prop Omitted	Prop Correct	Prop Omitted	Prop Correct	Prop Omitted
Addition									
Title I	80	.80	.03	.78	.04	.77	.04	.71	.06
Regular Education	40	.82	.02	.81	.02	.79	.05	.77	.04
Subtraction									
Title I	81	.74	.06	.69	.07	.66	.09	.61	.13
Regular Education	36	.76	.04	.68	.08	.72	.09	.65	.06
Multiplication									
Title I	83	.60	.10	.52	.13	.51	.15	.46	.21
Regular Education	43	.61	.09	.56	.13	.52	.15	.46	.19
Division									
Title I	74	.50	.23	.43	.24	.45	.21	.37	.27
Regular Education	38	.60	.18	.57	.21	.58	.23	.48	.20
Literal Comp.									
Title I	81	.70	.02	.57	.08	.63	.05	.47	.08
Regular Education	32	.81	.01	.65	.12	.73	.02	.66	.04

In Tables V-4 and V-5, it is apparent that performance was stable across the first three waves of testing for all cohorts. There is no evidence of overall growth or decline across time for either Title I or regular students. It is possible that this apparent stability masks other effects. For example, if the negative reaction seemingly exhibited in the fourth wave is present to some degree in earlier testings, then it may have neutralized a real increase in performance occurring simultaneously. On the other hand, it may be argued that the restrictiveness of the institutional environment is not conducive to learning and that in the absence of a strong instructional effect, performance may be expected to remain the same or even decline. Students may be doing well to maintain their level of performance over time. The data obtained in this study do not indicate which alternative is most reasonable: students may be gaining in achievement, but showing none of the gain due to reactions against the testing; they may be doing well by maintaining performance levels which would be expected to decline if education programs were not offered; they may be doing just as expected--maintaining a constant level of performance in an environment which is unlikely to change their performance levels. The last hypothesis requires the least explanation, but is the most discouraging.

The third result shown in Tables V-4 and V-5 is that students in programs of only regular education showed higher achievement at all waves of testing than students in Title I programs. Statistical tests were performed on the differences between Title I and regular education students at each point in time for each roster group. In arithmetic, 16 out of 20 comparisons were significant at $p < .05$; in literal comprehension, 19 out of 20 comparisons were significant at $p < .05$. These differences show that students assigned to Title I programs were on the average more in need of basic skills instruction than students assigned only to regular education instruction.

A fourth finding in Tables V-4 and V-5 is that Title I students entering at wave 2 seemed to perform worse than Title I students entering at other times, whereas regular education students entering at wave 2 performed better than their peers who entered at other times. One factor related to this widening difference is that the students assigned to Title I in that cohort were younger than

students assigned to Title I from all other cohorts, while students assigned only to regular education were older than those assigned only to regular education from the first and last cohorts. (The age of the third cohort was comparable.) One can only speculate as to why these age differences occurred in the second cohort. The assignment strategies used by institutions clearly tend to be biased toward placing younger students into Title I, as noted in Chapter IV. However, the reason for the increasing age gap is not clear. Perhaps institutions were able to make more careful (and more age-biased) assessments of need at that time because the populations of the institutions were not growing rapidly.

A variance components analysis was performed in order to estimate the components of variance on these measures for institutions and for pupils within institutions. In literal comprehension there was some tendency for the variance components due to students (between-student, within-institution) to increase slightly over time. This result indicates that the performance of students within institutions may have become more heterogeneous over time. Examination of the range of scores of the same cohort within institutions across waves of testing revealed that the highest score attained usually stayed about the same whereas the lowest score attained dropped markedly. This result suggests that decreasing motivation may have been operating at least among some students between waves 1 and 3.

The overall results for students from the randomly sampled institutions who had complete data in arithmetic and language arts at waves 1, 2, and 3 are displayed in Figure V-1. These results show that on the average Title I students started the testing phase at a lower level of achievement than regular education students. The performance of neither Title I nor regular students, however, showed much change over time. The slight increases from wave 1 to wave 3 shown in Figure V-1 are not statistically significant.*

---

*To investigate whether Title I instruction was effective in improving students' proficiency in basic skills, analyses of covariance were performed, controlling for initial achievement and other student characteristics. The results of these analyses, presented in Appendix H, show no differences between Title I and regular education. These analyses are presented in an appendix because we feel that, due to biased assignment of students to educational program, the analysis of covariance has questionable validity for these data.

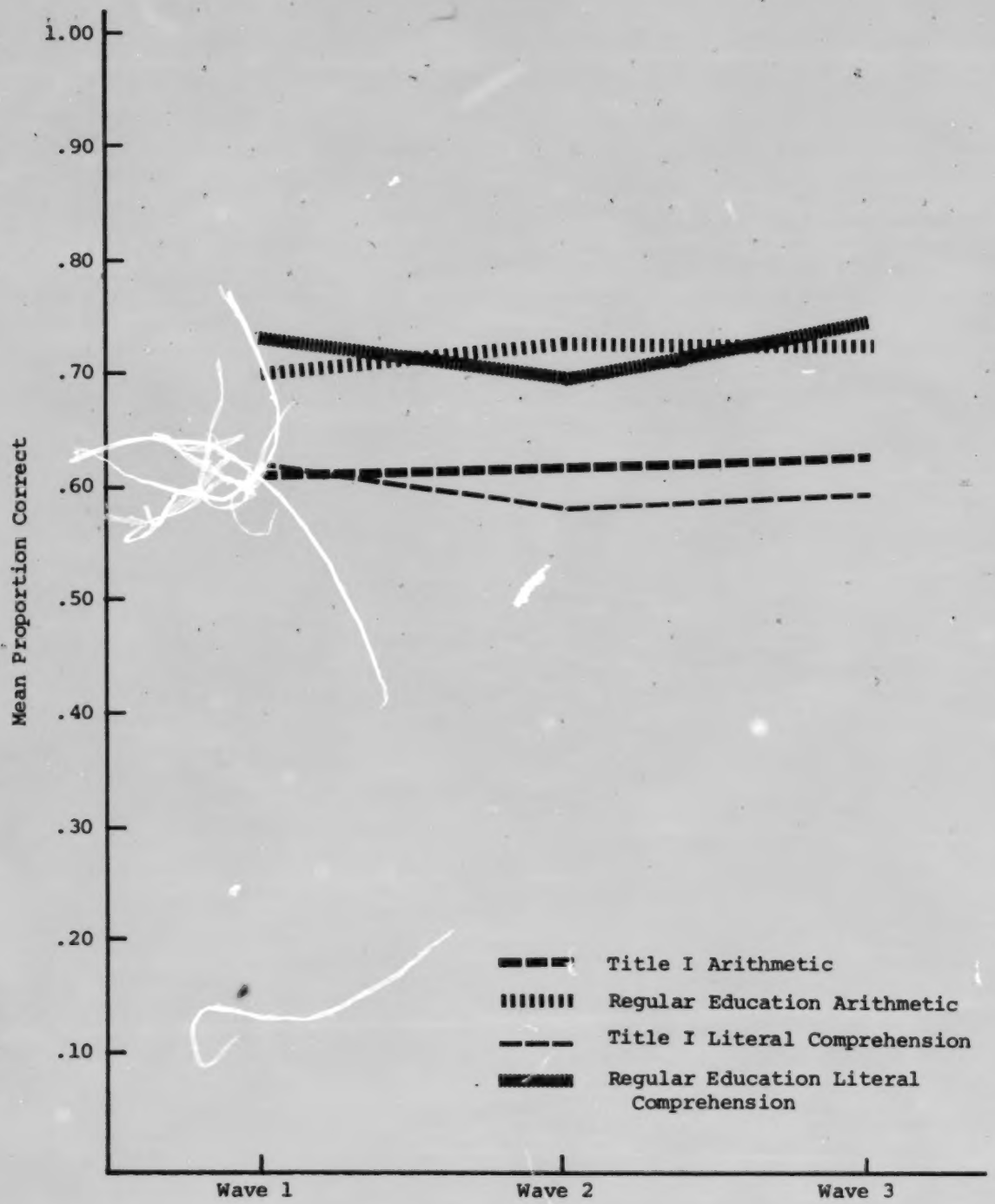

Figure V-1. Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I and Regular Education Students in Randomly Selected Institutions (Weighted by Selection Probability)

The results described above do not distinguish between institutions. The overall stability of performance masks variability between institutions. Appendix G gives the mean proportion correct on arithmetic and literal comprehension across waves of testing for each randomly selected institution. As can be seen in this appendix, some institutions showed increases in performance among Title I or regular education students whereas others showed decreases or mixed results.

These results show that educational programs in some institutions successfully promoted improvement over twelve weeks whereas educational programs in others did not. The results in a few institutions are displayed as graphs in Figure V-2 (a) to (d). These graphs represent the variety of performance patterns among institutions.

A striking finding in these graphs is the parallel movement of Title I and regular students within an institution in a subject matter domain. Figures V-2 (a) and (b) are good examples of this. The curves for Title I and regular education students in arithmetic are nearly parallel, as are the curves for literal comprehension, even though the curves for arithmetic have a different shape from those in literal comprehension. In other institutions the curves for arithmetic and literal comprehension are nearly parallel, as shown in Figures V-2 (c) and (d). This parallelism within institutions may be partly a function of the effects of collective testing. Some students in an institution may set an example for others to follow--expressing enthusiasm or apathy toward the testing situation. Parallel lines show some indication that Title I and regular educational programs were exerting the same amount of influence; Title I students generally did not improve relative to regular education students, nor did they deteriorate relative to them.

Detailed questions that may have been masked by these overall null results are: (1) Do students with low initial achievement improve over time? Or stated in another way, is Title I (or regular education) instruction effective for students with low initial achievement? and (2) Among those students who show improvement in performance over time, do Title I and regular education students show different amounts of growth? These questions will be addressed below.

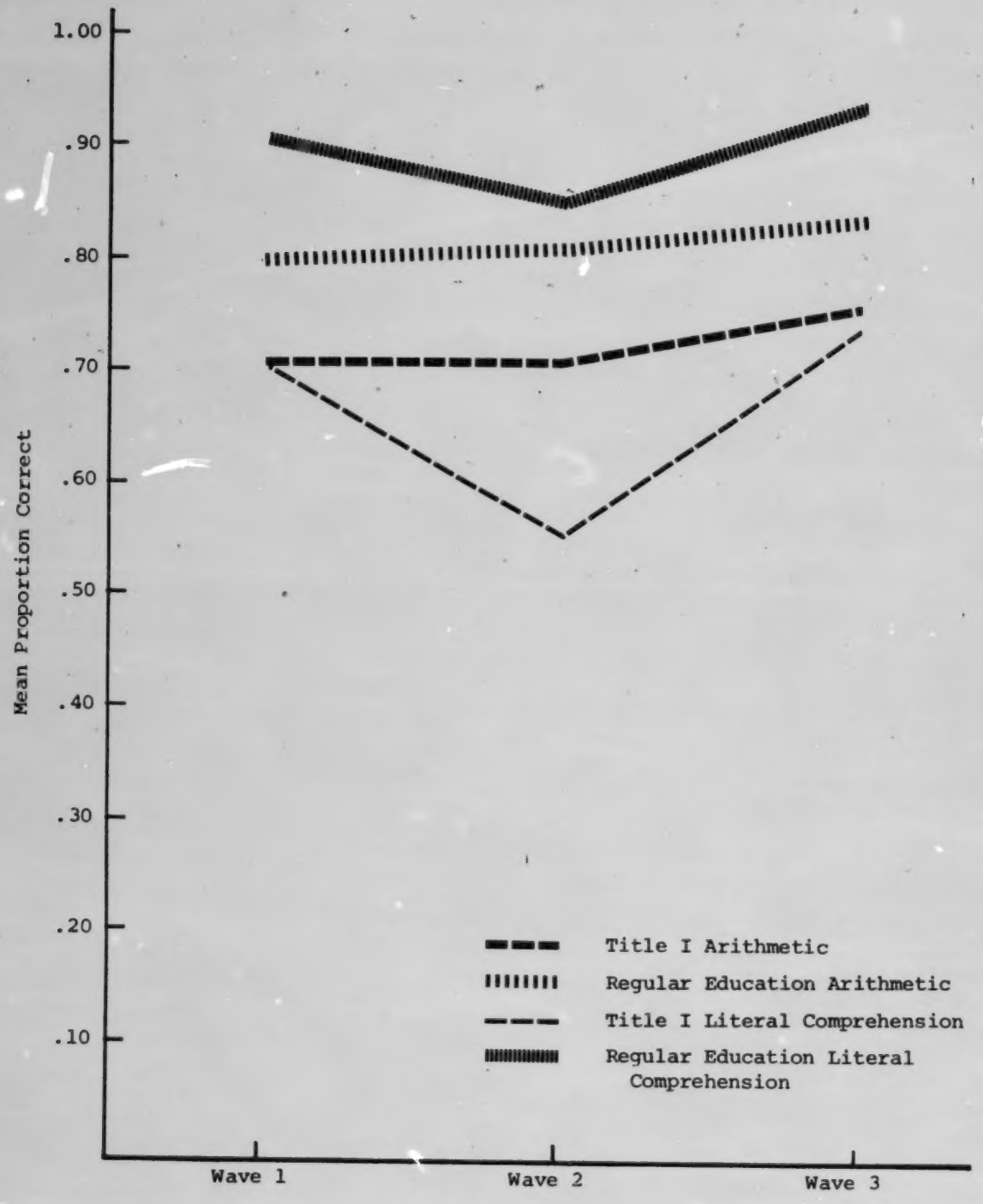

Figure V-2a. Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I and Regular Education Students at Institution 17

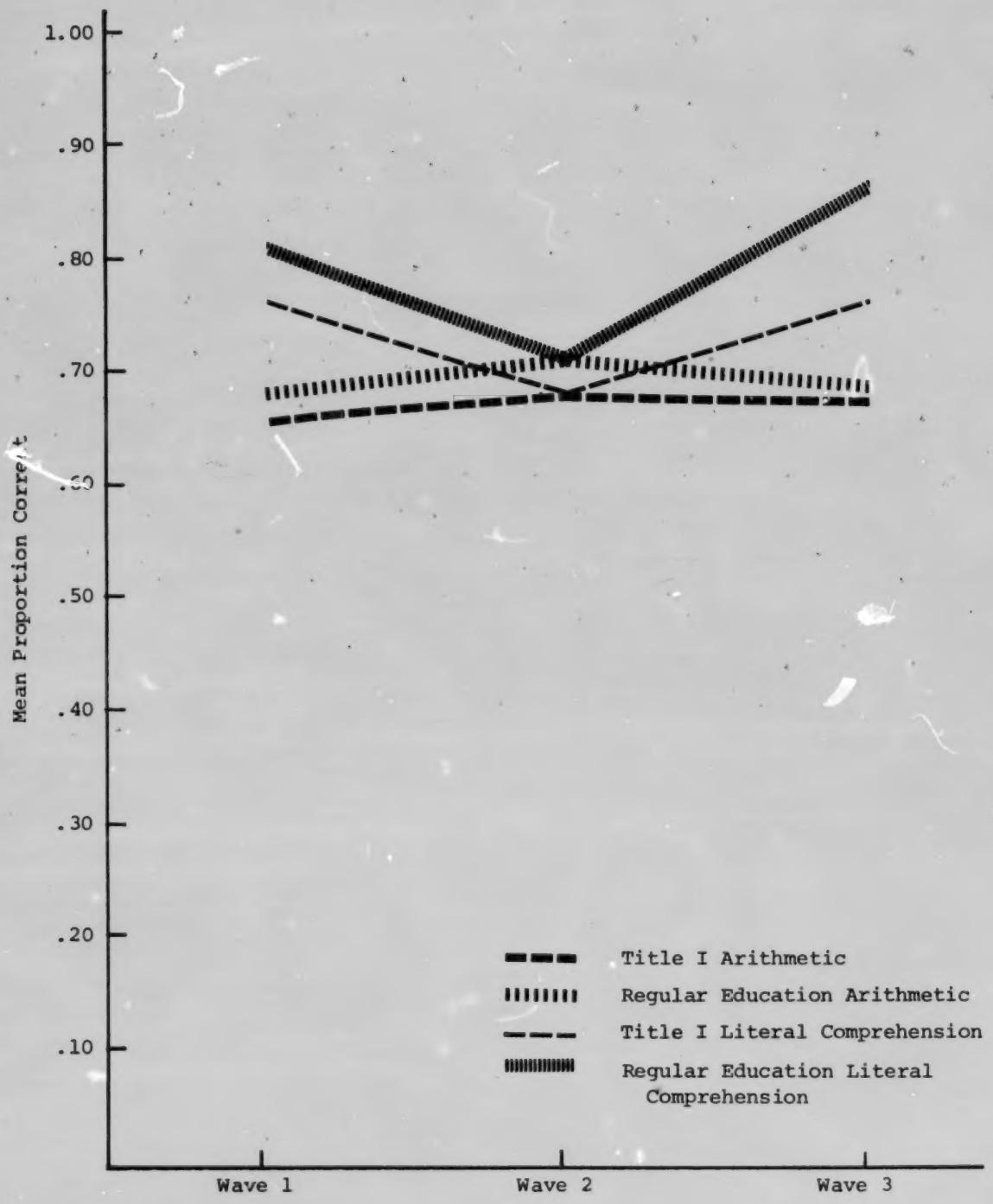

Figure V-2b. Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I and Regular Education Students at Institution 26

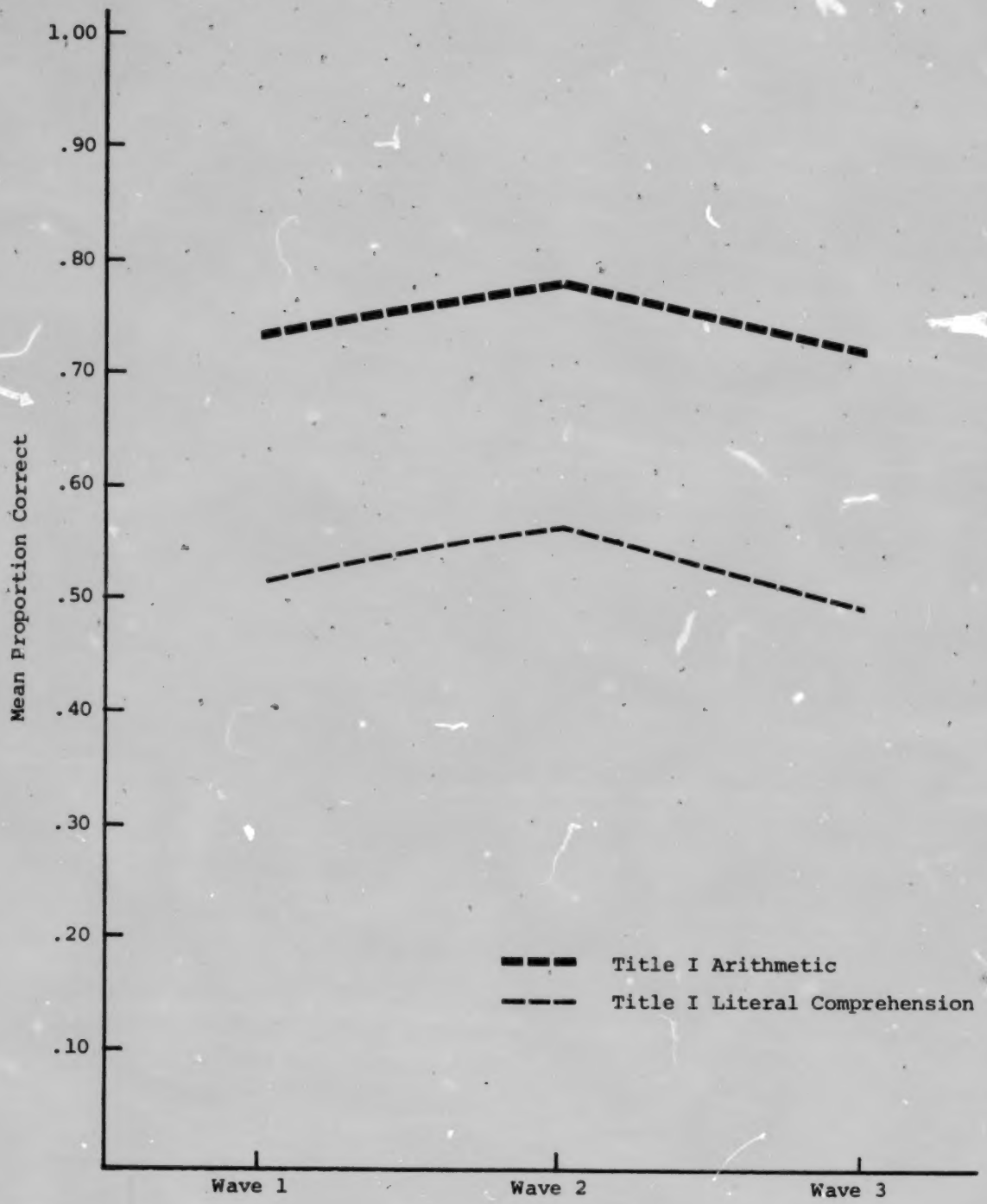

Figure V-2c. Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I Students at Institution 7

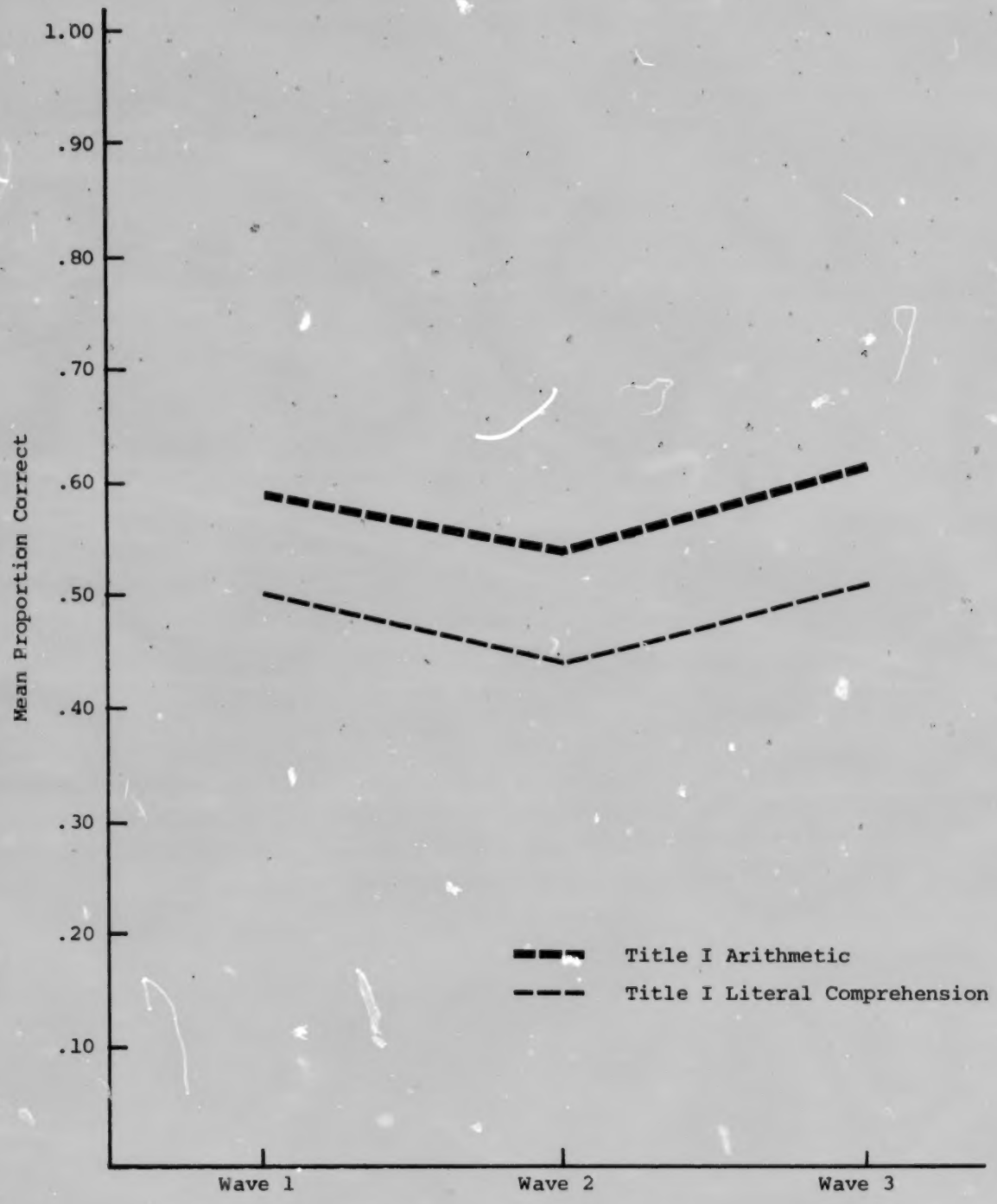

Figure V-2d. Mean Proportion Correct in Arithmetic and Literal Comprehension for Title I Students at Institution 16

To answer the first question, the performance of Title I and regular education students was examined separately in groups defined by wave 1 scores. Mean performance over time for each group is presented for arithmetic in Table V-7 and for literal comprehension in Table V-8. These tables show that students with low initial scores improved, while students with relatively high initial scores worsened over time. Regression effects due to unreliability in the tests could account for this result. That is, the scores of students who scored low initially may have underrepresented their true achievement. Upon retest their scores would tend to be closer to the mean of their group. The opposite effect would be expected for students scoring high initially. To examine whether the regression effect accounted for the results shown in Tables V-7 and V-8, the true scores for these students were estimated using data on the reliability of these tests.* The "true" performance by groups is presented in Tables V-9 and V-10. The means of estimated true scores still indicate that the low-achievers improved performance over time while higher-achievers worsened over time. Several explanations are possible. First, Title I and regular education instruction may have been more effective for students scoring low initially. Second, low-achievers may have experienced increasing motivation over time while higher-achievers experienced decreasing motivation. The data available are insufficient to distinguish among these causes.

Because scores that decrease over time may not be valid indicators of a student's achievement, the performance of students whose scores were maintained or improved over time was examined separately. The mean performance of those students is given in Table V-11. These results indicate that, for those students who maintained their level of performance or improved over time, Title I students showed about the same average gain as regular education students. The slight increase of Title I students relative to regular education students for language arts is not statistically significant.

---

*Although the internal consistency reliability (coefficient alpha) for arithmetic was .90 and the reliability for language arts was .93, we estimated that the test-retest reliability of these tests was probably about .75. This value was used to estimate "true scores".

Table V-7. Arithmetic Performance by Score Group and Program
(Score Groups Defined at Wave 1)

Score Group/Program	n	Initial Achievement	Final Achievement	Change From 1st to 3rd Testing
1 (0-25%)				
Title I	112	.44	.47	.03
Regular Education	24	.44	.46	.02
2 (25-50%)				
Title I	92	.61	.61	.00
Regular Education	52	.61	.61	.00
3 (50-75%)				
Title I	95	.75	.72	-.03
Regular Education	48	.75	.76	.01
4 (75-100%)				
Title I	82	.87	.84	-.03
Regular Education	66	.89	.89	.00

Table V-8. Language Arts Performance by Score Group and Program
(Score Groups Defined at Wave 1)

Score Group/Program	n	Initial Achievement	Final Achievement	Change From 1st to 3rd Testing
1 (0-25%)				
Title I	129	.26	.31	.05
Regular Education	31	.29	.46	.17
2 (25-50%)				
Title I	119	.63	.58	-.05
Regular Education	44	.64	.64	.00
3 (50-75%)				
Title I	110	.80	.75	-.05
Regular Education	50	.81	.77	-.04
4 (75-100%)				
Title I	80	.90	.84	-.06
Regular Education	83	.92	.88	-.04

Table V-9. Estimated "True"* Arithmetic Performance by Score Group and Program (Score Group Defined at Wave 1)

Score Group/Program	n	Initial Achievement	Final Achievement	Change From 1st to 3rd Testing
1 (0-25%)				
Title I	112	.50	.52	.02
Regular Education	24	.52	.54	.02
2 (25-50%)				
Title I	92	.62	.62	.00
Regular Education	52	.64	.64	.00
3 (50-75%)				
Title I	95	.72	.70	-.02
Regular Education	48	.74	.74	.00
4 (75-100%)				
Title I	82	.80	.78	-.02
Regular Education	66	.84	.84	.00

Table V-10. Estimated "True"* Language Arts Performance by Score Group and Program (Score Groups Defined at Wave 1)

Score Group/Program	n	Initial Achievement	Final Achievement	Change From 1st to 3rd Testing
1 (0-25%)				
Title I	129	.37	.39	.02
Regular Education	31	.42	.54	.12
2 (25-50%)				
Title I	119	.62	.58	-.04
Regular Education	44	.67	.67	.00
3 (50-75%)				
Title I	110	.75	.70	-.05
Regular Education	50	.79	.76	-.03
4 (75-100%)				
Title I	80	.81	.76	-.05
Regular Education	83	.86	.84	-.02

*True scores were estimated using a weighted average of the individual's score and the group mean.

Table V-11. Performance of Students Who Maintained or Improved Scores Over Time

Subject Matter	Program	n	Mean Score		Change From 1st to 3rd Testing
			Wave 1	Wave 3	
Arithmetic	Title I	135	.60	.67	.07
	Regular Education	71	.72	.79	.07
Language Arts	Title I	164	.56	.67	.11
	Regular Education	89	.71	.80	.09

ANALYSIS OF PURPOSIVELY SELECTED SITES

The ten purposively selected sites were compared to the randomly selected sites to ascertain whether differences in patterns of growth were demonstrated. It should be recalled that the purposive selection was based on Phase I data collectors' observations of program characteristics such as strong staff commitment to instruction, strong administrative leadership, student involvement in learning, positive staff-student interactions, etc. This analysis revealed no systematic differences between the purposively selected or randomly selected sites. Consequently, no separate documentation of the characteristics of purposively selected sites are reported. However, purposively selected sites are represented in a later section presenting the results of exploratory analyses that investigated the relationships of institutional and educational program characteristics to achievement.

STUDENT ATTITUDES

Tables V-12 through V-22 show the results of the attitude surveys given at the first and third waves of testing. The form of these tables is similar to that used for the DRT scores: each cohort of students is subdivided into groups by test-taking patterns.

Self Esteem and Locus of Control

Students tended to rate themselves highly on the self-esteem scale. These results contradict much of the literature which, based mostly on clinical assessments of institutionalized and delinquent youths, cites their negative self-images. However, in a previous study where a measure of self-esteem based on self-assessment methods was used (Bennett et al., 1971), similarly high self-esteem scores were obtained.

The group-testing method of obtaining the data, where the items were read to the students, may have contributed to a positive bias--students might be more inclined to display a positive self-image when they are in a group setting where there may be some element of competition or need to project positive impressions. This

Table V-12. Responses to Self Esteem Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-7)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	6.15	--	6.12	--
	s =	1.25	--	1.32	--
	n =	430	--	171	--
Waves 1 and 3	Mean =	6.01	6.20	5.94	6.17
	s =	1.46	1.32	1.58	1.45
	n =	308	308	120	120
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	5.68	--	6.08
	s =	--	1.54	--	1.23
	n =	--	1.29	--	65
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	6.09	--	6.29
	s =	--	1.51	--	1.17
	n =	--	248	--	136

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-13. Responses to Locus of Control Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-12)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	7.84	--	8.38	--
	s =	2.79	--	2.80	--
	n =	418	--	169	--
Waves 1 and 3	Mean =	7.45	7.79	8.07	8.31
	s =	3.14	3.00	3.05	2.96
	n =	284	284	122	122
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	7.08	--	8.59
	s =	--	2.72	--	2.42
	n =	--	119	--	57
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	7.66	--	9.20
	s =	--	2.91	--	2.56
	n =	--	233	--	127

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-14. Responses to Peer Affiliation Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-3)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	2.27	--	2.23	--
	s =	0.90	--	0.97	--
	n =	451	--	173	--
Waves 1 and 3	Mean =	2.22	2.19	2.06	2.26
	s =	0.98	0.94	.94	0.90
	n =	341	341	128	128
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	2.22	--	2.37
	s =	--	0.90	--	0.79
	n =	--	133	--	62
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	2.27	--	2.36
	s =	--	0.91	--	0.92
	n =	--	263	--	135

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-15. Responses to School Involvement Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-3)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	2.32	--	2.17	--
	s =	0.93	--	1.08	--
	n =	466	--	184	--
Waves 1 and 3	Mean =	2.29	2.19	2.23	2.36
	s =	1.01	1.00	1.02	0.90
	n =	350	350	138	138
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	2.44	--	2.19
	s =	--	0.82	--	1.03
	n =	--	134	--	68
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	2.50	--	2.62
	s =	--	0.82	--	0.71
	n =	--	265	--	145

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-16. Responses to Liking of School Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-5)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	2.66	--	2.99	--
	s =	1.59	--	1.72	--
	n =	459	--	181	--
Waves 1 and 3	Mean =	2.71	2.61	2.91	2.86
	s =	1.68	1.66	1.77	1.63
	n =	340	340	135	135
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	2.46	--	3.67
	s =	--	1.64	--	1.57
	n =	--	137	--	65
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	2.88	--	3.45
	s =	--	1.55	--	1.65
	n =	--	259	--	142

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-17. Responses to Sociability Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-4)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	3.30	--	3.16	--
	s =	1.01	--	1.19	--
	n =	420	--	177	--
Waves 1 and 3	Mean =	3.36	3.34	3.13	3.54
	s =	0.97	1.06	1.24	0.78
	n =	312	312	124	124
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	3.29	--	3.34
	s =	--	0.87	--	1.07
	n =	--	127	--	60
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	3.43	--	3.16
	s =	--	0.93	--	1.18
	n =	--	248	--	131

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-18. Responses to Attitudes Toward Mathematics Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-3)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	1.94	--	2.02	--
	s =	1.10	--	1.05	--
	n =	469	--	187	--
Waves 1 and 3	Mean =	1.99	1.89	1.99	2.00
	s =	1.10	1.14	1.10	1.12
	n =	352	352	144	144
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	1.70	--	1.63
	s =	--	1.18	--	1.05
	n =	--	141	--	69
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	1.75	--	1.95
	s =	--	1.19	--	1.12
	n =	--	270	--	151

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-19. Responses to Attitudes Toward Reading Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-3)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	2.20	--	2.32	--
	s =	0.93	--	0.90	--
	n =	468	--	186	--
Waves 1 and 3	Mean =	2.34	2.28	2.37	2.42
	s =	0.88	0.90	0.85	0.82
	n =	352	352	141	141
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	2.21	--	2.38
	s =	--	0.88	--	0.83
	n =	--	136	--	70
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	2.22	--	2.45
	s =	--	0.91	--	0.74
	n =	--	270	--	150

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-20. Responses to Teacher Support Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-7)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	4.93	--	4.80	--
	s =	1.96	--	1.96	--
	n =	440	--	1.73	--
Waves 1 and 3	Mean =	4.88	4.95	4.70	4.75
	s =	2.02	2.04	1.93	1.95
	n =	320	320	126	126
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	5.06	--	4.74
	s =	--	1.96	--	2.19
	n =	--	122	--	62
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	4.72	--	5.14
	s =	--	1.97	--	1.75
	n =	--	254	--	133

Means and standard deviations were estimated by weignting student data for probability of selection.

Table V-21. Responses to Practical Orientation of Learning Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-4)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	3.29	--	3.11	--
	s =	1.18	--	1.31	--
	n =	445	--	175	--
Waves 1 and 3	Mean =	3.19	3.12	3.07	3.15
	s =	1.28	1.29	1.20	1.20
	n =	314	314	129	129
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	3.20	--	3.27
	s =	--	1.27	--	1.32
	n =	--	127	--	63
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	3.10	--	3.38
	s =	--	1.32	--	1.12
	n =	--	262	--	136

Means and standard deviations were estimated by weighting student data for probability of selection.

Table V-22. Responses to Expectations for Release Success and Stigmatization Scale
(Data from Randomly Sampled Institutions)
(Score range: 0-4)

Results for:		Title I Plus Regular Education		Regular Education Only	
		Wave 1	Wave 3	Wave 1	Wave 3
Students First Tested at Wave 1					
Students Tested at:					
Wave 1 Only	Mean =	2.46	--	2.57	--
	s =	1.12	--	1.07	--
	n =	431	--	173	--
Waves 1 and 3	Mean =	2.40	2.62	2.50	2.52
	s =	1.06	1.13	1.19	1.13
	n =	316	316	125	125
Students First Tested at Wave 2					
Students Tested Only at Wave 3	Mean =	--	2.33	--	2.99
	s =	--	1.16	--	1.05
	n =	--	125	--	65
Students First Tested at Wave 3					
Students Tested Only at Wave 3	Mean =	--	2.47	--	2.75
	s =	--	1.02	--	1.09
	n =	--	250	--	131

Means and standard deviations were estimated by weighting student data for probability of selection.

interpretation of these relatively high scores is consistent with the literature on inmate subcultures, codes, and victimization which indicates that mechanisms such as projecting positive impressions are used to gain status in the institutional environment (Toch, 1977; Bartollas et al., 1976, Irwin, 1970). Conversely, institutionalized students may be more inclined to display a negative self-image to a counselor in a one-on-one setting--perhaps feeling that such self-deprecation will gain them sympathy. The true self-image of these youths is something of an unresolved mystery.

Responses to the 'locus-of-control' scale (which measures a person's perception of the likely effects of his or her actions on outcomes) were also positively skewed. Higher scores mean that a person's locus of control is more 'internal'; that is, the person believes that his or her actions do influence outcomes. Lower scores mean that the person feels that factors outside his or her control (fate, luck, powerful others) determine outcomes to a greater extent.

Student Attitudes Toward School

In general, students reported a high degree of involvement with learning tasks. Sixty percent responded positively to all scale items measuring school involvement. Greater variation existed on "liking of school", with 40 percent reporting a great dislike of school in general. While this measure may indicate negative feelings for schools as students have experienced them, it may not totally capture their capabilities or motivation to learn. For example, 96 percent of the students responded positively to the statement "I like to learn about new things". School-related tasks such as reading and math appeared to be favorably regarded. Over 40 percent of the students responded positively to all assessments of liking for math and reading tasks.

For the most part, students were not attending schools in the institution because they expected to be released earlier if they did. Over 60 percent of the students responded negatively to a statement that "doing well in school will help me get out of this place sooner". Seventy percent of the students reported that they planned to return to school upon their release.

Student Perceptions of the Educational Environment

Students provided assessments of the educational environment in terms of peer affiliation, teacher support, and practical orientation of the education program.

Students reported positive assessments of peer support within institution contexts, since 60 percent of students responded favorably to all measures of peer affiliation. Perceptions of teacher support were more varied: almost one-fourth of the students reported that their teachers were unsupportive or non-caring.

Students tended to agree that the learning that occurs in correctional education contexts has a practical orientation. Eightly percent of the students rated education programs high on this dimension.

While students in Title I programs differ from those in regular education programs, an analysis-of-covariance indicated that there were no differential rates of change for the two groups. These analyses are reported in Appendix H.

THE RELATIONSHIP OF CHARACTERISTICS OF EDUCATIONAL PROGRAMS TO ACHIEVEMENT AND ATTITUDE MEASURES

This section will report the results of analyses directed towards the detection of the relationships among the characteristics of educational programs and achievement and attitude measures. Before presenting these results, a brief description of the analytical process is given.

Analytical Process

The analyses presented in the following pages should be considered more exploratory than confirmatory. While there were expectations that each of the characteristics of institutions and educational programs which were reported in the data collection could influence either achievement scores or attitudes, there were no strong theories which specifically forecast confirmation.

Scores of variables were processed in performing these analyses. All of the characteristics of institutions and educational programs described in Chapter IV were included, as well as certain additional variables. In these analyses

all 40 institutions were used because the purpose of the analysis was to discover relationships rather than to estimate population characteristics. Thus, if some purposively selected institutions had especially effective programs, and differed from other institutions in terms of program characteristics, then these facts would be detected in the data analysis.

All student variables were aggregated to the institutional level because the institution is the appropriate level at which to conduct these analyses. Analyses were performed separately for aggregates of students in Title I plus regular instruction and for students in programs of regular instruction only. Only those students who participated in the tests at both wave 1 and wave 3 were included in these aggregates.

The first analytical step after creating the aggregated, institutional-level data base was to regress a wave 3 score on a wave 1 score and obtain the residuals, e.g., that part of the wave 3 score not accounted for by the wave 1 score. In the second step, these residuals were then correlated with all of the characteristics of institutions and programs (including aggregated student background measures, such as age). The largest 25 correlations were determined from each of these analyses.

In the third step, the lists of variables having high correlations with the residuals were combined into one list. Redundancies were eliminated in this list and the remaining variables were included with the aggregated wave 1 scores in building new prediction models for the wave 3 aggregate scores. Some variables, such as the per pupil costs, were forced into the analysis at this step because of their intrinsic policy interest, even though they had not been selected in the earlier exploratory data analysis steps.

The variables which made up a prediction model were tried in various combinations in order to select the most parsimonious model consistent with the data. The usual strategies of stepwise regression, or permuted subset analysis, were not employed because patterns of missing data made it impossible to fit models with more than eight or nine variables with any confidence.

Generally speaking these models involve only a few variables and seem to account for a reasonable proportion of the variance in the outcome measures, although certain of the relationships may be due to chance. A few of the results do not seem readily interpretable, and these will be noted in the discussion that follows. The first testing score is always included in the model even though it is not statistically significant in two analyses. This ensures that the estimated partial regression weight of each of the other variables in each of the models is properly interpreted as that variable's influence on the change in the outcome measure from the first testing to the third testing.

Results of Relational Analyses - Achievement Scores

Tables V-23 to V-26 show the results of these analyses for the achievement scores of Title I students and regular education only students in both reading and mathematics.

These tables show that very few program characteristics were significantly related to achievement when the wave 1 test score was included in the model. The characteristics which did enter the various models seem to be related to the outcomes in the expected directions. Thus, the size of the Title I reading program is negatively related to reading achievement gains for both Title I students and those in programs of only regular education. Larger programs seem to be more detrimental than smaller programs. This effect is rather small, requiring an additional 100 students to cause a drop of 4 percentage points in the scores of Title I students (or a drop of 10 percentage points in the scores of regular education only students).

The Educational Program Administrator's ranking of the importance to the institution of custody goals versus other institutional objectives is negatively related to the reading achievement gains of students in programs of only regular education instruction. A gain in rank of one point (on a scale of 0 to 6) is associated with a loss of 1.6 percent correct among these institutions, for students not in Title I. Other variables in this ranking (e.g., academic education goals, counseling goals) did not relate to achievement outcomes. Apparently the custody ranking distinguishes institutions emphasizing incarceration from those

Table V-23. The Relationship of Educational Program Characteristics to the Proportion Correct on the Literal Comprehension DRT at Wave 3 for Students in Programs of Title I Plus Regular Education Instruction

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Wave 1 Score	0.90	.08	.0001
Number of Title I Reading Students	-.00045	.00019	.025

R^2 = .77: Mean square error = 0.062. (Model based on 38 observations with complete data.)

Table V-24. The Relationship of Educational Program Characteristics to the Proportion Correct on the Literal Comprehension DRT at Wave 3 for Students in Programs of Only Regular Education Instruction

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Wave 1 Score	0.85	.09	.0001
Number of Title I Reading Students	-0.0010	.00025	.0003
EPA Ranking of Custody Goals Versus Other Institutional Goals	-0.016	.0061	.015

R^2 = .85: Mean square error = 0.018. (Model based on 28 observations with complete data.)

Table V-25. The Relationship of Educational Program Characteristics to the Proportion Correct on the Arithmetic DRT at Wave 3 for Students in Programs of Title I Plus Regular Education Instruction

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Wave 1 Score	1.001	.088	.0001
R^2 = .78: Mean square error = 0.027. (Model based on 38 observations with complete data.)			

Note: None of the educational program characteristics were related to the wave 3 test scores when the wave 1 score was entered into the equation.

Table V-26. The Relationship of Educational Program Characteristics to the Proportion Correct on the Arithmetic DRT at Wave 3 for Students in Programs of Only Regular Education Instruction

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Wave 1 Score	0.948	0.089	.0001
EPA Encourages Innovation	0.037	0.018	.05
Facility Director Estimate of Education Staff Turnover	-.0030	0.0011	.02
R^2 = .89: Mean square error = 0.012. (Model based on 24 observations with complete data.)			

emphasizing other aspects of the institution, e.g., academic education, vocational education, counseling, institutional upkeep.

None of the characteristics of institutions or education programs was related to the wave 3 scores of the Title I students on the arithmetic DRT once the wave 1 score was entered into the model. Among the institutional aggregates of students in programs of only regular education, two additional variables appeared to be significant: the extent to which the EPA encourages teachers to be innovative within their own courses of instruction is positively related to arithmetic gain scores, while the estimate of the percentage of educational staff turnover is negatively related to these scores.

At the risk of considerable over-simplication and over-generalization, it seems reasonable to assert that smaller programs, with low staff turnover, when directed by an EPA who encourages innovation, and when lodged in institutions where custody is a secondary concern, will generally be more successful.

This set of relational analyses may not have been sensitive to the important characteristics of education programs for two reasons: 1) the short time span of 12 weeks between wave 1 and wave 3 may not have been long enough to allow some effects to become manifest, and 2) the small size of the sample may make some of the effects hard to detect. These problems are interrelated in that the longer the time span, the more likely it is that a particular institution would have its first wave sample exhausted. If there are very good short-term programs, it may be very hard to detect their effects.

The stability of the institutional-level average scores over the 12-week period of instruction should, perhaps, alert one to the possibility that instructional interventions may not be as powerful as is sometimes hoped. Juveniles who are performing at a 5th-grade or lower level have successfully resisted instructional effects for several years. Their defenses may be stronger than the instructional programs, and may minimize the effects that the provision of substantial time, effort, and money could have.

Results of Relational Analyses - Attitudinal Measures

The eleven attitude measures which were derived from the student attitude survey, given on the first and third visits to the institutions, were analyzed for their relationship to measures of program and institutional characteristics. Formally, this analysis proceeded in the same fashion as the corresponding analyses for the achievement measures. Institution-level aggregates (means) for each testing were created for each of the eleven attitude measures in each of the two programs--regular education only, and Title I plus regular education. This led to 22 analyses in which a third-testing attitude score (average) was regressed upon a first-testing score (average) together with variables representing the characteristics of programs and institutions. An initial choice of 25 variables for each such analysis was made by selecting those variables in the large set which correlated most highly with residuals defined by regressing the third-testing score on the first-testing score. This set of variables was further winnowed down by performing repeated multiple regression analyses, successively deleting variables which did not reach a .05 level of significance. Any predictor variable (characteristic of program or institution) that survived these analyses was tried in all the other models to determine if it could contribute to an understanding of the other variables as well. The winnowing process was performed again and the final results are presented in Tables V-27 to V-48.

As happened in the analyses of achievement outcomes, there are fewer institutions with complete data for non-Title I outcomes than there are for Title I outcomes. The samples of non-Title I students simply shrank very rapidly in the 12 weeks between the first and third testing, leaving 11 institutions with no matched cases to include in the process of constructing averages.

Self-Esteem

The self-esteem scale results (Tables V-27 and V-28) are not very consistent. The major element of consistency seems to be that the percentage turnover in non-educational treatment staff (during the previous year) is negatively related to change in self-esteem. This could mean that higher proportions of new

Table V-27. Regression Analysis for Average Self-Esteem Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Self-esteem (first testing)	0.594	0.143	.001
Percentage turnover in the non-educational treatment staff (during the previous year)	-0.010	0.004	.022
EPA ranking of the importance of reading instruction in the Title I program (higher ranks mean lesser importance)	-0.417	0.159	.016
Proportion of education programs which are accredited.	-0.225	0.081	.012

R^2 = .70: Mean square error = .92. (Model based on 25 observations with complete data.)

Table V-28. Regression Analysis for Average Self-Esteem Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Self-esteem (first testing)	0.225	0.130	.092
EPA rating of quality of Title I furniture	0.252	0.108	.026
Percentage turnover in the non-educational treatment staff (during the previous year)	-0.009	0.003	.008
EPA self-rating of support for the Title I program	-0.454	0.190	.024

R^2 = .56: Mean square error - 1.65. (Model based on 35 observations with complete data.)

Table V-29. Regression Analysis for Average Locus of Control Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Locus of control (first testing)	0.831	0.111	.001
EPA ranking of the importance of academic learning within the institution	0.656	0.251	.015

R^2 = .71: Mean square error = 6.09. (Model based on 28 observations with complete data.)

Table V-30. Regression Analysis for Average Locus of Control Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Locus of control (first testing)	0.860	0.065	.001
Title I funds spent for educational administrators (in thousands of dollars)	-0.053	0.020	.014
EPA perception of degree to which Title I program should be run by the teachers	0.282	0.117	.023
EPA does not agree that for some students basic skills education can be set aside until they are ready	1.168	0.229	.001

R^2 = .88: Mean square error = 3.75. (Model based on 34 observations with complete data.)

Table V-31. Regression Analysis for Average Attitude Toward Reading Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Attitude toward reading (first testing)	0.527	0.077	.001
Composite EPA rating of the tendency of institutional personnel other than teachers and the EPA to make decisions about educational programs	-0.578	0.105	.001
Number of FTE teacher aides in academic education programs	-0.640	0.020	.006
Title I expenditures on items other than staff in programs other than reading or mathematics (in thousands of dollars)	-0.018	0.006	.006

R^2 = .85: Mean square error = .22. (Model based on 22 observations with complete data.)

Table V-32. Regression Analysis for Average Attitude Toward Reading Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Attitude toward reading (first testing)	0.484	0.092	.001

R^2 = .43: Mean square error = 0.740. (Model based on 39 observations with complete data.)

Table V-33. Regression Analysis for Average Attitude Toward Mathematics Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Attitude toward mathematics (first testing)	0.569	0.118	.001
EPA ranking of the preferred importance of vocational education	0.284	0.076	.002
EPA indication that low student motivation is not a problem	0.452	0.116	.001

R^2 = .73: Mean square error = .68. (Model based on 22 observations with complete data.)

Table V-34. Regression Analysis for Average Attitude Toward Mathematics Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Attitude toward mathematics (first testing)	0.558	0.088	.001
Number of FTE remedial specialists in mathematics in the academic education program	0.313	0.084	.001
Number of months for which regular education teachers are salaried	0.142	0.027	.001
EPA composite rating of the lack of conflict between education and other institutional priorities.	0.102	0.024	.001

R^2 = .75: Mean square error = 0.54. (Model based on 35 observations with complete data.)

Table V-35. Regression Analysis for Average Affiliation Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Affiliation (first testing)	0.555	0.196	.010
Funding for education from sources other than Title I N or D (in hundred-thousands of dollars)	-0.105	0.047	.035

R^2 = .46: Mean square error = 1.11. (Model based on 25 observations with complete data.)

Table V-36. Regression Analysis for Average Affiliation Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Affiliation (first testing)	0.657	0.120	.001

R^2 = .45: Mean square error = 0.91. (Model based on 39 observations with complete data.)

Table V-37. Regression Analysis for Average School Involvement Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
School involvement (first testing)	0.645	0.099	.001
EPA indicator that ungraded classes are used extensively in the regular education program	0.347	0.111	.005

R^2 = .64: Mean square error = 0.543. (Model based on 28 observations with complete data.)

Table V-38. Regression Analysis for Average School Involvement Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
School involvement (first testing)	0.524	0.108	.001
EPA indication that low student motivation is not a problem	0.182	0.057	.003
Proportion of education programs which are accredited	-0.090	0.039	.026
Percentage of institutional residents who are neglected youth	0.008	0.003	.010

R^2 = .61: Mean square error = .62. (Model based on 38 observations with complete data.)

Table V-39. Regression Analysis for Average Liking of School Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Liking of school (first testing)	0.857	0.151	.001

R^2 = .55: Mean square error = 3.35. (Model based on 28 observations with complete data.)

Table V-40. Regression Analysis for Average Liking of School Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Liking of school (first testing)	0.563	0.079	.001
Observer ratings of the degree to which Title I program is not distinct from the regular education program	0.269	0.127	.041

R^2 = .61: Mean square error = 1.57. (Model based on 39 observations with complete data.)

Table V-41. Regression Analysis for Sociability Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Sociability (first testing)	0.896	0.177	.001
Percentage turnover in the non-educational staff (during the previous year)	-0.008	0.003	.040

R^2 = .59: Mean square error = .84. (Model based on 25 observations with complete data.)

Table V-42. Regression Analysis for Average Sociability Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Sociability (first testing)	0.299	0.146	.049
EPA indicator that ungraded classes are used extensively in the regular education program	0.360	0.110	.003
Non-Title I expenditures on buildings (in thousands of dollars)	-0.042	0.015	.008

R^2 = .49: Mean square error = .79. (Model based on 36 observations with complete data.)

Table V-43. Regression Analysis for Average Teacher Supportiveness Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Teacher supportiveness (first testing)	0.519	0.146	.002
EPA rating of the cooperativeness between students and teachers	1.044	0.259	.001

R^2 = .59: Mean square error = 3.14. (Model based on 27 observations with complete data.)

Table V-44. Regression Analysis for Average Teacher Supportiveness Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Teacher supportiveness (first testing)	0.512	0.139	.001
Students are selected for educational programs through the recommendations of the institutional staff	-0.973	0.242	.001

R^2 = .50: Mean square error = 5.23. (Model based on 39 observations with complete data.)

Table V-45. Regression Analysis for Average Practical Orientation of Learning Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Practical orientation of learning (first testing)	0.820	0.128	.001

R^2 = .61: Mean square error = 1.22. (Model based on 28 observations with complete data.)

Table V-46. Regression Analysis for Average Practical Orientation of Learning Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Practical orientation of learning (first testing)	0.382	0.159	.022
Title I funds spent for educational administrators (in thousands of dollars)	-0.042	0.015	.010

R^2 = .30: Mean square error = 2.35. (Model based on 36 observations with complete data.)

Table V-47. Regression Analysis for Average Post-Release Expectation and Stigmatization Scores of Non-Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Post-release expectations and stigmatization (first testing)	0.317	0.154	.053
Funding for education from sources other than Title I N or D (in hundred-thousands of dollars)	-0.227	0.051	.001

R^2 = .54: Mean square error = 1.33. (Model based on 25 observations with complete data.)

Table V-48. Regression Analysis for Average Post-Release Expectation and Stigmatization Scores of Title I Students at the Third Testing

Source	Estimated Partial Regression Coefficient	Standard Error	p <
Post-release expectations and stigmatization (first testing)	0.699	0.080	.001
EPA indicator that students are assigned to regular education classes on the basis of achievement	-0.229	0.066	.002
Observers ranking of the importance of therapy as an institutional activity	0.061	0.024	.016

R^2 = .79: Mean square error = .52. (Model based on 39 observations with complete data.)

treatment staff lead to this negative outcome or that other features of the institution contribute to higher staff turnover and declining self-esteem scores simultaneously.

The first-testing score is strongly related to the third-testing score for Title I students only. This does not have a ready explanation. Nor do the negative relationships found between gains in self-esteem and the proportion of education programs which are accredited, or the Educational Program Administrator rating of his or her own support for the Title I program. However, it may be comforting to know that the more highly ranked that reading stands among activities in the Title I program, the larger the gain in self-esteem, but a direct interpretation of this relationship is difficult.

Locus of Control

The locus-of-control scale (a measure of one's self-perceived ability or inability to influence outcomes) resulted in models which seem more tractable (Tables V-29 and V-30). Clearly, the first measure is highly related to the second (taken at the third testing) for students in both programs. For non-Title I students, the higher the institution's ranking of academic learning, the larger the gain of locus of control, that is, the more students changed to the perception of themselves as in control of their own destiny. For Title I students, the expenditure of Title I funds on educational administrators was related to students feeling less in control of their own destinies, while the extent to which the EPA felt the Title I program should be developed and run by teachers, and the EPA indicator that instruction in basic skills should not wait upon students becoming "ready", are both positively associated with students' feeling more in control. These results may suggest that institutions with over-bureaucratized educational programs perhaps make students feel less in control of what happens to them than may be true of less-bureaucratized programs.

Attitudes Toward Reading

The results of the relational analyses for attitudes toward reading (Tables V-31 and V-32) are interesting for the paucity of variables which seem to influence

the average scores of Title I students. For non-Title I students the tendency to have decisions about the educational program made by people not connected directly to that program, the number of FTE teacher aides in academic education programs (including, but not limited to Title I), and the Title I expenditures on items other than staff in programs other than reading or mathematics are all negatively related to changes in average attitudes toward reading. These relationships may be indicative of the ways in which academic education programs may be slighted by institutions--by allowing non-education staff to control the programs, by hiring aides rather than teachers, and by buying equipment and supplies for use in non-basic skills programs.

Attitudes Toward Mathematics

The analyses reported in Tables V-33 and V-34 are an interesting contrast to those resulting from the analysis of changes in attitudes toward reading. The variables which entered the models for attitudes toward mathematics are all positively related to gains in attitudes toward mathematics and seem to describe the things that could be "right" about running an educational program. For the Title I students these variables are: (1) the number of remedial specialists in mathematics, which reflects the commitment to mathematics instruction; (2) the length of the employment year for regular education teachers, which indicates a fuller commitment to education (through 12-month rather than 9-month years); and (3) harmony between educational activities and other institutional priorities. For the non-Title I students the absence of problems with student motivation is also positively related to gains in attitudes toward mathematics, as is the preferred ranking of vocational education. The latter may reflect an institutional desire to provide educational programs which emphasize practical applications of a computation and/or measurement.

Affiliation

The relational analyses did not prove very fruitful for the measure of affiliation (the extent to which students work and interact with other students in the educational setting). Other than the first-testing score on this measure,

only the funding for education from sources other than Title I seemed to be related to gains, and this relationship was negative and weak (Tables V-35 and V-36).

School Involvement

With one exception, the variables which appear in models for the school involvement measure (the extent to which students participate in the daily functioning of educational activities) seem reasonably related to the gains in this measure (Tables V-37 and V-38). For non-Title I students, the EPA indicator that ungraded classes are used extensively in the regular education program is positively related to gains in school involvement. Classrooms which contain mixtures of grade levels may be run in a fashion which encourages student participation and involvement--perhaps by being more individualized. For Title I students, the positive association of the absence of student motivational problems with gains in school involvement seems very reasonable--though the causal direction is not clear. The percentage of institutional residents who are neglected youth relates positively to increasing involvement because in institutions serving this population, education is the main focus of the day's activities. The negative relationship of the proportion of education programs which are accredited seems to be unexplainable.

Liking of School

Here again, the relational analyses did not discover much of interest (Tables V-39 and V-40). The field staff that collected data for this study were requested to indicate whether the Title I program seemed to be distinct from the regular education program in the institutions they visited. The analysis shows that the less distinct the two programs the more Title I students gain in their liking for school. They may feel less isolated or less set apart in settings where the two programs are indistinct. This may deserve more attention in any further studies of these programs because federal policy emphasizes the maintenance of a distinction. Other characteristics of institutions may simultaneously influence the lack of distinction between programs and the increased liking for school. These would need to be explored also.

Sociability

Models for the 'sociability' mean scores at the third testing are presented in Tables V-41 and V-42. For non-Title I students, the percentage turnover in non-educational treatment staff during the previous year is negatively related to gains in average sociability scores. For Title I students, the extensive use of ungraded classes is positively related to gains in average sociability scores, while the spending of non-Title I money on buildings is negatively related to such gains. The percentage turnover in non-educational treatment staff also appeared in both the Title I and non-Title I models for self-esteem scores. This finding may indicate that there is a tendency for institutions with high turnover rates to be less interested in promoting the self-esteem and sociability of their students, or that both staff turnover and unsociability may result from unmeasured institutional problems. The use of ungraded classes would probably tend to break down what are typically age-specific groups than would the use of graded classes, while new-plant expenditures may reflect a diminished attention to the problems of current students being served at the expense of long-range objectives (e.g., the practice of spending sought after funds for permanent resources rather than for the expendable resources such as staff salaries).

Teacher Supportiveness

Models for the average 'teacher supportiveness' scores at the third testing are presented in Tables V-43 and V-44. The EPA rating of cooperativeness between students and teachers is positively related to gains in the average teacher supportiveness scores for non-Title I students. Some institutions may be able to create a cooperative atmosphere where student feelings that they have the support of their teachers grow over time.

For Title I students, the selection of students for educational programs on the basis of institutional staff recommendations is associated with decreases in the average teacher supportiveness scores. It would be interesting to know how this relationship arises. Perhaps the institutions which use personal recommendations tend to elicit feelings that some residents are more favored than others, and that teachers are not generally supportive.

Practical Orientation of Learning

The models for the average scores on practical orientation of learning at the third testing (Tables V-45 and V-46) yield only one variable associated with gains: the expenditure of Title I funds on educational administrators was negatively related to gain scores for Title I students. This may mean that institutions which spend Title I funds for administrators rather than instructional services (e.g., more staff or more varied program offerings) tend to have Title I programs which are viewed by Title I students as less relevant to intended post-release activities. It is tempting to conclude that this is an effect of the size of the institutions--where larger institutions might use more Title I money to hire administrators and, simultaneously, be more likely to run Title I programs which seem less relevant to the students. However, direct measures of the size of the program could have appeared in the equation, but did not. The effect seems to be associated with the amount of Title I money spent for administrators regardless of the size of the institution or the Title I program.

Post-Release Expectations and Stigmatization

Models for the mean scores on 'post-release expectations and stigmatization' at the third testing are presented in Tables V-47 and V-48. For the non-Title I students, the amount of funding for education from sources other than Title I N or D funds is negatively related to gains in this measure. This probably means that institutions with larger education budgets tend to serve residents who feel more and more that they will not be able to succeed after release and that they will suffer stigmatization for having been in an institution.

For the Title I students, gains in post-release expectations are negatively related to the assignment of students to regular education classes based on achievement, and positively related to the relative importance of therapy as an institutional activity. Perhaps segregating students on the basis of achievement encourages Title I students to think less of their post-release chances for success. An emphasis on therapy could mean paying specific attention to the encouragement of students to think positively about their expectations for post-release success.

It should be re-emphasized that the models reported in Tables V-23 through V-48 were developed not from a strong theoretical perspective, but by allowing the data to indicate which relationships were to be included, based on the relative predictive power of the variables describing institutions and educational programs. These analyses should not be seen as confirming the relationships which are described in the tables. They should serve, rather, as indicators which, upon interpretation, will yield testable hypotheses for subsequent studies.

The data show some consistency across models in that any program or institution variable which is positively related to gains on one outcome measure is also positively related to gains on any other outcome measure for which it is included in the model. This implies that the data reveal no conflicts between Title I and non-Title I programs in which changing an institutional characteristic would be good for one program and bad for the other. Generally, the relationships seem to indicate that activities which detract from education, or from attention to the attitudinal consequences of institutionalization, result in decreasing scores on the achievement and attitude outcomes measured in this study.

CONCLUSIONS AND RECOMMENDATIONS

The results reported in this evaluation of the impact of Title I education programs in state institutions for neglected or delinquent youth clearly indicate that students in Title I supplementary programs and those in programs of regular (state-supported) education only are seriously deficient in the basic skill areas of reading and arithmetic. The group of students served by Title I programs is somewhat more deficient than the group served by regular education only. However, about 30 percent of the students who are not receiving Title I programs is somewhat more deficient than the group served by regular student in these institutions. In response to this widespread need, about one-half of the institutions in the sample indicated that they provide Title I services to <u>all</u> eligible students. Some of the other institutions would like to serve more students. Clearly, there is a sizable group of students who could be served through expansion of the Title I program.

The results of receiving Title I services are less clear, however. Students receiving Title I supplementary services did not appear to gain or lose in basic skills proficiency over a twelve-week period of instruction. Similarly, students in programs of regular education only neither gained nor lost basic skills proficiency over this time period. While the Title I program does not appear to differentially affect outcomes, it may be accomplishing some success for the students it serves because they might not otherwise be given any educational services or might be unable to benefit from the services offered in the regular education program.

While the lack of a clear impact of Title I programs is surely disappointing, especially in view of the apparent concern instructional staff show for very individualized educational programs, the small effect may be due to the limited amount of exposure to instruction which students in state institutions receive. The data presented in this report indicate that the average student could receive between four and five hours per week of instruction in reading basic skills and from three to seven hours per week of instruction in arithmetic basic skills. The actual amounts of instructional exposure are even less: on the average, students actually receive from three to five hours per week of reading instruction and from two to two-and one-half hours per week of arithmetic instruction. Students in programs of regular education only receive more than one hour more of reading instruction per week than the students who are receiving the Title I supplementary services, while the students in Title I programs receive about one-half hour more instruction per week in arithmetic. Three sources of the large gap between the scheduled hours of instruction and the hours of instruction actually received may be identified:

- The institution interferes directly with the educational program of groups of students. This is most notable in Title I reading classes where, on the average, nearly thirty minutes per week of the scheduled classes were not held. Although the reason for cancelling classes was not ascertained, the likely causes are teacher illness, interference of other institutional programs (field trips, general work release, etc.), and lock-up of groups of students (which prevents them from receiving educational services).

- The institution interferes with individual students. This occurs when the institution schedules other activities (work release, counseling, parole hearings, etc.) in conflict with a student's instructional time. It also occurs when an individual student is punished by lock-up which prevents receiving educational services.
- Individual students do not attend classes for personal reasons. This occurs when students are sick, have gone AWOL from the institution, or simply do not attend the classes when they are held.

The last two effects cannot be sorted out in the data reported here. However, their combined effect is very powerful. Students in programs of regular education only lost 35 minutes of instruction per week in reading, on the average, due to these causes. Students in Title I programs in mathematics only attended one-third of the Title I mathematics instruction that was given; the average loss was two and one-half hours per week. Students in Title I programs in reading lost an average of 20 to 40 minutes per week of Title I supplementary classes in reading for these reasons. Clearly, the scheduling of classes and the assurance of attendance in the classes is an area which could be greatly improved in most institutions. Further information on the factors which inhibit the delivery and receipt of educational services (especially Title I services) will be presented in the report on the Substudy of Effective Practices, which is currently in progress.

Although there were no overall patterns of growth or change in either the attitudinal or basic skills achievement outcomes measured in this study, there were institution-to-institution differences in these changes. Many analyses were performed in an attempt to relate the differing characteristics of educational programs and institutions to the variation in changes in outcomes.

In general, smaller programs in institutions where education is considered to be an important activity, where the educational staff is able to run the program with little interference from the rest of the institution, and where there is little educational staff turnover seem to be most successful. Even in institutions with favorable standing on each of these characteristics, the small

amounts of instructional time combined with the short duration of stay of most residents (median residence is 7.5 months) raises serious questions about the ability of these institutions either to improve basic skills performance or to alter attitudes significantly. Furthermore, little effort seems to be expended to provide continuity for the students when they are released and would be expected to continue their education in non-institutional schools. More information on this aspect of the institutional programs will be presented in the report entitled Post-Release Experiences of Students From State Correctional Institutions (Volume 3 of this series) which was conducted as part of the study reported in this volume.

Before detailed policy recommendations can be drawn from the findings of the analyses presented in this report, at least two additional pieces of information are needed:

- A sense of how much influence federal and state policy changes could have on the characteristics of institutions and the educational programs within them.
- A more precise specification of the causal mechanisms by which characteristics of institutions and educational programs influence changes in achievement and attitudes.

The latter would provide a better estimate of the expected effects on achievement and attitude outcomes that could be realized by changing the institutions and the educational programs. Thus, the ultimate effect of policy on outcomes, via changes in institutions and educational programs, could be estimated.

While the analyses in the present report should not be treated as providing definitive estimates of the effects of institutional or educational program characteristics on outcomes, they do indicate some general concerns which need to be carefully considered in the development of educational programs for these institutions. Paramount among these, especially for the delinquent population, is the conflict between the intent of incarceration in an institution and the goals of education. Institutions for delinquent youth are not hospitable environments for education. They may not be actively hostile toward education,

but educational goals are not well integrated, in many settings, into the overall institutional goals. The findings of this study that badly needed training in basic skills is offered for very small amounts of time and that institutions seem to reorganize student schedules in ways which diminish the attendance in these classes, imply that basic skills training has a low priority in these institutions.

Many of the youth in these institutions are young enough to be subject to state regulations mandating their continued education. It appears that within the institutions they may not be given access to educational experiences which are suitable for their needs. Viewed either as the substantial core of any program of rehabilitation, or simply as an obligation of the state to all young persons, the provision of educational services in state institutions for neglected or delinquent youth should be treated as more important, relative to other institutional goals, than it has been thus far.

REFERENCES

Anderson, T.H. "Cloze Measures as Indices of Achievement Comprehension When Learning from Extended Prose." Journal of Educational Measurement, 1974, 11, 83-92.

Bartell, T. et al., Compensatory Education and Confined Youth: A National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth. Volume 1. TM-5792/005/00. Santa Monica, CA, System Development Corporation, 1977.

Bartollas, C., Miller, S. & Dinitz, S. Juvenile Victimization: The Institutional Paradox. New York: John Wiley and Sons, 1976.

Bennett, L., Sorensen, D., & Forshay, H. The Application of Self-Esteem Measures in a Correctional Setting. Journal of Research in Crime and Delinquency, 1971, 8, 1-9.

Bock, R.D. Multivariate Statistical Methods in Behavioral Research. New York: McGraw-Hill, 1975.

Bock, R.D. "Multivariate Analysis of Variance of Repeated Measurements" in C.W. Harris (Ed.) Problems in Measuring Change. Madison: University of Wisconsin Press, 1963, 85-103.

Bormuth, J.R. "Comparable Cloze and Multiple-Choice Comprehension Test Scores." Journal of Reading, 1967, 10, 291-299.

Bratten, J., Robertson, D., and Geddes, C. Final Report Job Corps Mathematics Program Revision Project. Santa Monica: System Development Corporation, 1972.

Carroll, J., Davies, P., and Richman, B. The American Heritage Word Frequency Book. New York: Houghton-Mifflin, 1971.

Carroll, J.B., and Freedle, R.O., eds. Language Comprehension and the Acquisition of Knowledge. Washington, D.C.: Winston, 1972.

Carver, R.P. "Measuring the Primary Effect of Reading: Reading-Storage Technique, Understanding Judgments, and Cloze." Journal of Reading Behavior, 1974, 6, 249-274.

Cronbach, L.J., Rogosa, D., Price, G. & Floden, R. Analysis of Covariance--Temptress and Deluder? Or Angel of Salvation? Stanford Evaluation Consortium, Stanford University, 1976.

Education Commission of the States, "An Overview of Findings and Recommendations of Major Research Studies and National Commissions Concerning Education of Offenders", 1976.

Florida Department of Education, A Guide to 1977-78 Statewide Assessment Results, (undated).

Glaser, R. & Nitko, A.J. "Measurement in Learning and Instruction." In Thorndike, R.L., ed. Educational Measurement. Washington, D.C.: American Council on Education, 1971.

Hafner, L.E. "Relationships of Various Measures to the Cloze." In Thirteenth Yearbook of the National Reading Conference. Milwaukee, Wisconsin: National Reading Conference. 135-145, 1964.

Harris, A., and Jacobson, M. Basic Elementary Reading Vocabularies. New York: Macmillan, 1972.

Hayes, J.R., ed. Cognition and the Development of Language. New York: Wiley and Sons, 1970.

Horton, R.J. "The Construct Validity of Cloze Procedures: An Exploratory Factor Analysis of Close, Paragraph Reading, and Structure of Intellect." Dissertation Abstracts International, 1973, 34, 3140.

Irwin, John. The Felon. Englewood Cliffs: Prentice Hall, 1970.

Kenoyer, C. Teacher Judgement of Need for Compensatory Education. Working paper, System Development Corporation, Santa Monica, CA, 1978.

Marchbanks, G. & Levin, H. "Cues by Which Children Recognize Words." Journal of Educational Psychology, 1965, 56, 57-61.

Millman, J. "Criterion-Referenced Measurement." In Popham, W.J., ed. Evaluation in Education: Current Applications. Berkeley, California: McCutchan Publishing Corporation, 1974.

McNeill, D. The Acquisition of Language: The Study of Developmental Psycholinguistics. New York: Harper and Row, 1970.

Moos, R. Evaluating Correctional Settings. New York: John Wiley and Sons, 1975.

Nowicki, S. & Strickland, B. A Locus of Control Scale for Children. Journal of Consulting and Clinical Psychology, 1973, 40, 148-154.

Ohnmact, F.W., Weaver, W.W., & Kohler, E.T. "Cloze and Closure: A Factorial Study." Journal of Psychology, 1970, 74, 205-217.

O'Reilly, R.P., Schuder, R.T., & Kidder, S.J. Validation of a Multiple-Choice Cloze Test of Literal Comprehension: Summary Report. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976.

197

O'Reilly, R.P. & Streeter, R.E. Reports on the Development and Validation of a System for Measuring Literal Comprehension in a Multiple-Choice Close Format: Preliminary Factor Analytic Results. Paper presented at the annual meeting of the Americal Educational Research Association, San Francisco, 1976.

Prange, J.L. "An Investigation of the Relationships Obtaining Between Cloze Test Measures of Reading Performance and Measures of Critical Reading, General Reading, Intelligence and Sex." Dissertation Abstracts International, 1974, 34, 6459.

Reagen, M. & Stoughten, D. School Behind Bars: A Descriptive Overview of Correctional Education in the American Prison System. Metuchen, N.J.: The Scarecrow Press, Inc., 1976.

Spencer, M., Bessemer, D., Fedan, N. & Offutt, B. Measures of Growth in Practical Achievement. Chapter III of Measures of Student Growth for the Sustaining Effects Study, TM-5693/003/00. System Development Corporation, Santa Monica, 1976.

Stevenson, C.R. & Hoover, H.O. 1977-78 Technical Report, Statewide Assessment Program, Florida Department of Education (undated).

Tallmadge, G.K. & Wood, C.T. User's Guide ESEA Title I Evaluation and Reporting System, RMC Research Corporation, Mountain View, California, 1978. (Available from Office of Evaluation and Diseemination, Office of Education U.S. Department of Health, Education and Welfare.)

Taylor, S., Frackenpohl, H., & White, C. A Revised Core Vocabulary: A Basic Vocabulary for Grades 1-8, An Advanced Vocabulary for Grades 9-13 (Research and Information Bulletin No. 5). Huntington, New York: Educational Developmental Laboratories, March 1969.

Toch, H. Living in Prison. New York: The Free Press, 1977.

Trickett, E. & Moos, R. Social Environment of Junior High School Classrooms. Journal of Educational Psychology, 1971, 65, 93-102.

Weaver, W.W., & Kingston, A.J. "A Factor Analysis of the Cloze Procedure and Other Measures of Reading and Language Ability." Journal of Communications. 13, 252-261, 1963.

Webb, N. & Keesling, J.W. Handbook for Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth, TM-5792/005/00. Santa Monica: System Development Corporation, 1978.

Wenk, E. & Moos, R. Social Climates in Prison: An Attempt to Conceptualize and Measure Environmental Factors in Total Institutions. Journal of Research in Crime and Delinquency, 1972, 9, 134-148.

APPENDIX A

DEVELOPMENT OF THE DOMAIN-REFERENCED TESTS

DOMAIN IDENTIFICATION AND DEFINITION

In order to develop domain-referenced tests in the basic skills area (language arts and mathematics), it was necessary first to identify a comprehensive set of possible domains. This was followed by a choice of the most important domains for testing in the Phase II evaluation. This section discusses the procedure used to arrive at the final set of domains.

The initial step was a survey of the content of existing commercially available tests, with particular attention to criterion-referenced tests. Although these tests were not domain-referenced, their content was typically organized in terms of a conceptual content hierarchy, thus allowing the identification of content domains. A total of 40 criterion-referenced tests, covering language arts, reading, and math for grades K-12, was identified and reviewed. In addition, four test publishers provided extensive catalogues of behavioral or instructional objectives, which were also reviewed.

Finally, information was obtained from a number of research reports describing in detail the structure (in terms of objectives and content) of published reading and mathematics curricula. These reports were particularly valuable because they reviewed published texts that are widely used in public school systems.

The results of the review indicated extensive agreement on the content areas (domains) represented under the general rubric of "basic skills". A comprehensive list of domains was developed from these sources, including 16 language arts/reading domains and 17 mathematics domains. The domains are described in Tables A-1 and A-2.

The next step was to choose the most important domains for testing in Phase II. The list of domains was sent to each member of the Policy and Research Advisory Groups. They were asked to rank the domains in order of their importance to Title I programs, separately for language arts/reading and for

APPENDIX A

Table A-1. Description of Language Arts/Reading Domain

Domain	Description
Discrimination	Identify visual and auditory language symbols.
Phonetic Analysis	Blend letter sounds in order to read words from a basic vocabulary.
Structural Analysis	Identify the structural elements of words (base words, compounds, contractions, syllables, affixes).
Word Recognition	Identify words that belong to a basic vocabulary.
Vocabulary	Understand the meaning of words that are part of a basic vocabulary.
Literal Comprehension	Understand the meaning of material that has been read.
Interpretive Comprehension	Interpret, summarize, or infer meaning from material that has been read.
Critical Comprehension	Identify missing, misleading or irrelevant information in material that has been read, differentiate fact vs. opinion, logical vs. emotional appeals or identify bias in material that has been read.
Spelling	Use letter sound correspondence to correctly spell words from a basic vocabulary.

Table A-1. Description of Language Arts/Reading Domain (Cont'd)

Domain	Description
Punctuation and Capitalization	Perform punctuation or capitalization on written passages so that it is grammatically correct.
Language Usage	Identify parts of speech; identify correct and incorrect usage of various parts of speech in written passages.
Syntax	Identify and apply correct sentence structure for various types of sentences; identify functions of words and phrases in context and existence of syntax errors in written passages.
Locating Materials	Use library and other references sources to locate information required by a specific task.
Organization	Identify purpose of reading/studying; plan studying to encompass time constraints and purposes; categorize and sequence information into outline form.
Pictorial and Graphic Skills	Read and interpret maps, graphs, tables.
Following Directions	Follow and correctly execute a series of directions for accomplishing a specific task.

Table A-2. Description of Mathematics Domains

Domain	Description
Pre-Mathematics	Identify objects that differ in amount, size, shape, color, or numerosity.
Addition of Integer Numbers	Solve problems involving the addition of integer numbers.
Subtraction of Integer Numbers	Solve problems involving the subtraction of integer numbers.
Multiplication of Integer Numbers	Solve problems involving the multiplication of integer numbers.
Division of Integer Numbers	Solve problems involving the division of integer numbers.
Fractions, Decimals, Mixed Numbers	Compare fractions, decimals and mixed numbers; solve problems involving the addition, subtraction, multiplication, or division of decimals, fractions and mixed numbers.
Number Principles	Understand number type, positional notation, identities, inverse operations, and the commutative, associative and distributive properties of numbers.
Equations, Inequalities, Word Problems	Analyze equations, inequalities, word problems and number sentences; solve word problems through use of equations, number sentences, reflexive, symmetric or transitive properties.
Ratio, Proportion, Per Cent	Identify and compare ratios; convert fractions, proportions and per cents; solve problems involving the use of proportions, ratios, and per cents.
Practical Mathematics	Use mathematical knowledge and skills to solve common practical problems.
Sets	Understand and use set notation and set membership; perform operations involving sets.

Table A-2. Description of Mathematics Domains (Cont'd)

Domain	Description
Geometry	Identify points, angles, lines, and plane figures; perform operations involving these; understand concept of symmetric congruence, intersection.
Coordinate Geometry	Identify points in a coordinate system, and write the ordered pairs naming the points; plot simple functions.
Probability and Statistics	Interpret graphs and tables, perform computation of basic statistics for tabled data; identify the probability associated with specific events.
Measurement	Compare and convert measurements of time, money, weight length or capacity; solve problems involving such measurements.
Exponents, Roots, Expanded Notation	Identify scientific, expanded exponential notation, perform operations involving roots and exponents.
Logical Thinking	Identify the truth or falsity of compound statements involving connectives; answer questions concerning the sequence or steps or the decision arising from a flowchart.

mathematics. The mean rankings for each group are presented in Tables A-3 and A-4. It was felt that the rankings given by the Policy Advisory Group (PAG) should be given the most weight because the PAG members were most familiar with actual Title I N or D programs.* It is interesting to note that these rankings compare favorably with the amount of time spent on related activities as reported in the results of the Phase I analyses (Bartell, et al., 1977). This offers some evidence of correspondence between the panel's perception of instructional importance and the actual instructional emphasis of Title I programs.

The amount of time available for testing during the visits to each site also had to be weighted in determining the final selection of domains. Because the language arts subtests would probably be the most time consuming, only the three domains ranked in the top quartile of importance were chosen: word recognition, vocabulary, literal comprehension.

The mathematics domains could be tested more rapidly than the language arts domains, so more domains were included in the mathematics area: addition, subtraction, multiplication, division, and practical mathematics. Pre-mathematics was eliminated because the trivial content might have proven insulting to the students. The content of the four basic operations domains was extended to include decimals, common fractions, and mixed numbers. (Most of the problems used whole numbers.)

The following sections of this Appendix trace the development of the tests through precise domain specification to test generation. The language arts tests are presented first, followed by the mathematics tests.

---

*This weighting is reflected in the Total column of Tables A-3 and A-4. The weighted rankings given in the Total column were used to order the table entries from most important (top) to least important (bottom).

Table A-3. Language Arts/Reading Domain Rankings

Domain	Rankings by Advisory Groups			Time Spent by Teachers*
	RAG	PAG	Total	
Vocabulary	4.0	3.9	3.9	2.6
Word Recognition	5.0	3.3	4.0	2.4
Literal Comprehension	2.6	5.4	4.2	2.6
Discrimination	8.8	4.4	6.2	2.0
Spelling	7.4	5.7	6.4	2.2
Following Directions	7.6	5.7	6.5	2.6
Phonetic Analysis	8.4	6.3	7.2	2.3
Interpretive Comprehension	6.4	8.7	7.8	2.3
Punctuation, Captialization	9.2	9.3	9.2	N/A
Structural Analysis	11.4	9.3	10.2	2.0
Language Usage	10.0	12.4	11.4	N/A
Critical Comprehension	8.0	14.0	11.5	2.0
Pictorial, Graphic Skills	10.0	12.7	11.6	1.9
Locating Materials	11.6	12.3	12.0	2.0
Syntax	9.8	15.4	13.1	N/A
Organization	12.2	14.1	13.3	N/A

*Item Scale: 1 - Little time
2 - Some time
3 - Much time

Table A-4. Mathematics Domain Rankings

Domain	Rankings by Advisory Groups			Time Spent by Teachers*
	RAG	PAG	Total	
Addition	1.8	2.9	2.5	2.8
Subtraction	2.8	3.6	3.3	2.8
Multiplication	3.8	4.4	4.2	2.8
Pre-Mathematics	7.2	3.1	4.7	N/A
Division	4.8	5.1	5.0	2.8
Practical Math	5.2	6.0	5.7	2.4
Fractions, Decimals, Mixed	7.0	6.4	6.6	N/A
Measurement	9.6	8.6	9.0	2.2
Number Principles	9.0	9.8	9.5	1.9
Ratio, Proportion, Percent	8.8	10.5	9.8	N/A
Sets	11.0	12.0	11.6	1.5
Equations, Word Problems	12.3	12.6	12.5	2.1
Geometry	12.4	13.9	13.3	1.5
Logical Thinking	10.8	15.4	13.8	N/A
Coordinate Geometry	15.0	14.8	14.8	N/A
Probability, Statistics	13.8	16.4	15.5	N/A
Exponents, Roots	15.8	15.5	15.6	1.9

*Item Scale: 1 - Little time
2 - Some time
3 - Much time

DEVELOPMENT OF THE LANGUAGE ARTS TESTS

Domain Specification

In order to establish the content area (the format of the items and criteria for distractors), a review of relevant research literature and test instruments was undertaken. Combining this information with the constraints imposed by the testing situation in the institutional setting led to the selection of specific item types and the rules for generating items to be used on the tests.

The review of appropriate content and item types was focused on existing criterion-referenced tests. The different item types found in these tests were compiled into a separate list for each domain. These lists are presented in Tables A-5 to A-7. (Not all tests contained all of the item types; a few item types were fairly common across tests, while other item types were found only on one or two tests.)

The constraints on testing that were considered included: the number of areas to be tested, the respondent burden associated with other tests and student questionnaires, the attention span of the students, and the students' ability to deal with complex instructions. In discussions with the PAG, it was determined that the total testing time for reading should be limited to one to one and one-half hours, instructions should be simple, and the tasks should be as straightforward as possible. The number of different item types used was limited in order to avoid confusion. Further, because these are domain-referenced tests, the item types had to be amenable to clearly-specified, objective item generation techniques. These considerations provided the basis for the selection of item types and content elements.

The item type chosen for word recognition was one in which the student selected the response word matching the stimulus word. The other item types involved extraneous information, were irrelevant to practical word recognition applications, or unsuitable for item-generation techniques. The content of word recognition was individual nouns, verbs, adjectives, and adverbs chosen from a basic

Table A-5. Possible Item Types for Measurement of Word Recognition Skills

1. The student selects a given word from visually similar words in a sentence.
2. The student selects the word which names a picture when the choices are visually similar.
3. The student selects (from a list of words) the word that is the same as the stimulus word.
4. Given words all identical but one, the student will identify the one that is different.
5. Given a word and four variations of that word (e.g., backwards, upside-down), the student will identify the word that is identical to the stimulus word.

Table A-6. Possible Item Types for Measurement or Vocabulary Skills

1. The student selects the synonym (antonym, homonym) for the stimulus word.
2. The student selects the pair of words that are synonyms (antonyms, homonyms).
3. The student selects the synonym (antonym) for an underlined word in a sentence.
4. The student completes a word analogy.
5. The student selects the word that has the appropriate meaning to make sense in an incomplete sentence.
6. The student selects the picture that illustrates the word (or vice versa).
7. The student will identify the meaning of symbols, abbreviations and figurative expressions.

Table A-7. Possible Item Types for Measurement of Literal Comprehension Skills

1. The student selects the picture described by a sentence or paragraph (or vice versa).
2. The student answers questions about details in a story, including characters, character traits, places, times, events, sequence of events, and causes.
3. The student selects the question that is answered by an underlined phrase (who, what, where, etc.)
4. The student selects a paraphrase for a given sentence.
5. The student selects a summary sentence for a given reading selection.
6. The student fills in blanks created in a passage by random deletions of words. (Cloze Procedure)

vocabulary list organized by grade levels. Such a word list had been compiled by the New York State Education Department* (O'Reilly, Schuder and Kidder, 1976) from Harris and Jacobson's Basic Elementary Reading Vocabularies (1972), EDL Research and Information Bulletin 5: A Revised Core Vocabulary (Taylor, Frackenpohl, and White, 1969), and The American Heritage Word Frequency Book (Carroll, Davies, and Richman, 1971). The list contains approximately 12,000 words ranging from Kindergarten to grade 13. Distractors for word recognition were visually similar, real words from the vocabulary list. (The item-generation algorithm further required that distractors begin with the same initial letter, be of approximately the same length, and be of the same part of speech as the word being tested. These characteristics were identified as the most important attributes for word recognition by Marchbanks and Levin, 1965.

In vocabulary (Table A-6) the most straightforward item type (finding a synonym for the stimulus word among a list of distractors) was chosen. As in word recognition, the item content for the vocabulary test was derived from the word list compiled by the New York State Department of Education. Both the stimulus words and distractors were chosen from that list. The synonyms (the correct answers) were chosen with the aid of Roget's Thesaurus, but such synonyms also had to be present on the word list.

An enormous amount of work, theoretical and experimental, has been done in the area of language comprehension, e.g., Carroll and Freedle, 1972; McNeil, 1970; Hays, 1970. Of particular interest here is the research on literal comprehension. A number of studies have attempted to establish the relationship between standardized tests of ability and achievement, specific multiple-choice reading tests, structure-of-intellect tests, and various forms of the cloze test procedure, e.g., Anderson, 1974; Bormuth, 1967; Carver, 1974; Hafner, 1964; Horton, 1973; Ohnmacht, Weaver, and Kohler, 1970; Prange, 1974; and Weaver and Kingston, 1963. While the results of these investigations are not totally consistent

---

*We would like to express our grateful thanks to the New York State Department of Education, to Director of Research and Evaluation, Carl E. Wedekind, and to Ms. Ruth Salter for their cooperation and assistance in providing these materials and information.

and compelling, there is some evidence that multiple-choice reading tests (and at least some of the cloze procedures) are capable of measuring literal comprehension as distinct from general verbal ability.

The New York State Department of Education had already developed reading comprehension tests based on a modified cloze technique (see O'Reilly and Streeter, 1976; O'Reilly, Schuder and Kidder, 1976). Analyses of these tests indicated that the particular cloze technique being used was a valid measure of literal comprehension. In addition, the tests were truly domain-referenced in that the items were generated objectively from a large domain of reading passages representative of the reading that students were assigned. The domain consisted of approximately 1500 reading passages, ranging in difficulty from grade 1 to college-level, each with a maximum length of 80 words.* The passages were randomly selected from textbooks in reading, language arts, social studies, science, and math; from newspapers and newsmagazines, and from consumer and reference materials. Three to 10 test items were generated for each passage by deleting approximately every fifth noun, verb, adjective, or adverb, and providing a list of words (the deleted word plus several distractors chosen from the word list discussed above) from which the student was to choose the correct word. Distractors for the deleted word were randomly selected from the word list with the condition that they be the same part of speech. It was decided to use these materials in the construction of the test of literal comprehension. Table A-8 gives the content/item-type specifications for the language arts domains.

After choosing the item types and content to be used, the remaining task was to specify how the actual items would be constructed. For each domain, specifications were written which identified (1) the specific objectives to be measured, (2) the instructions to be given for each item type, (3) the format of the item, (4) detailed steps for generating items, and (5) the number of items to be generated.

---

*We would like to thank Dr. Robert O'Reilly for his technical advice and assistance in the use of these passages.

Table A-8. Content Specifications for Language Arts/Reading Domains

Domain	Elements of the Domain	Item Types
Word Recognition	All words on word list from New York State Department of Education	Pick stimulus word from four listed words
Vocabulary	All words on word list obtained from New York State Department of Education	Pick synonym of stimulus word from four listed words
Literal Comprehension	1500 reading passages obtained from New York State Department of Education	Multiple-choice Cloze Procedure

The word lists from the New York State Department of Education used in the construction of word recognition and vocabulary items were stratified by grade levels. Words were chosen from these strata in proportions reflecting the presumed distribution of achievement among N or D students. Table A-9 shows the composition of these two tests.

For the composition of the literal comprehension subtest, the 1500 passages in the domain were classified by content and by reading difficulty level. Table A-10 shows the relationship between reading difficulty level and the grade level of the source materials. Passages were chosen with difficulty levels between 4 and 17, inclusive. Thus, passages from first grade materials were eliminated, and the probability that items would be drawn from grades 9 and 10 was reduced.

Table A-11 shows the distribution of passages across content areas by difficulty levels. The mathematics and science content areas were eliminated on the basis that they would be less relevant to the educational goals of most institutions. The passages in the language arts area were also eliminated on the grounds that discussions of grammar and punctuation would be too abstract for most institutionalized youth. Such passages often used examples in which the content was deliberately nonsensical in order to make the grammatical point more forcefully. The Policy Advisory Group advised that such passages would be highly inappropriate for this population of students. Because there were so few newspaper and newsmagazine passages, these were combined to form one content area.

Noting (from Table A-10) that the difficulty levels represented different spreads of readability scores, eight difficulty strata were created in the following way:

1) Levels 4, 5, and 6 each formed one stratum.

2) Subsequent strata were formed by combining adjacent difficulty levels:

   a) 7 was paired with 8.

   b) 9 was paired with 10.

   c) 11 was paired with 12.

Table A-9. Item Sampling for Word Recognition and Vocabulary Subtests

Domain	Total Number of Items	Item Sampling Procedure
Word Recognition	16	2 words from grades 1-2 4 words from grades 3-4 6 words from grades 5-8 2 words from grades 9-10 2 words from grades 11-12
Vocabulary	24	3 from grades 1-2 6 from grades 3-4 9 from grades 5-8 3 from grades 9-10 3 from grades 11-12

Table A-10. Distribution by Difficulty Level of Passages from Graded Reading/Literature Texts

Difficulty Level	Readability Range*	Grade Level of Source									
		1	2	3	4	5	6	7	8	9	10
1	1.0-1.4	13									
2	1.5-1.9	33	7	1							
3	2.0-2.4	9	20	5							
4	2.5-2.9		36	4	2	1	2				
5	3.0-3.4		1	12	2	4		1		1	
6	3.5-3.9			13	6	1	1				
7	4.50-4.74			6	16	7	3	1	2	2	1
8	4.75-4.99			2	13	5	8	5	4	5	2
9	5.00-5.24			2	13	12	10	7	6	1	2
10	5.25-5.49			1	8	9	6	4	3	2	4
11	5.50-5.74			2	7	6	13	2	5	7	2
12	5.75-5.99			1	3	2	6	3	5	5	3
13	6.00-6.24			1	6	2	9	15	2	6	2
14	6.25-6.49				3	2	4	6	2	7	2
15	6.50-6.74					2	1	5	10	4	1
16	6.75-6.99					1	3	5	10	5	1
17	7.00-7.24				1		2	3	7	7	5
18	7.25-7.49					1	1	4	3	12	7
19	7.50-7.74				1			1	2	4	5
20	7.75-7.99								1	7	10
21	8.00-8.24				1		1			2	5
22	8.25-8.49					1			2	2	8
23	8.50-8.74							1		1	
24	8.75-8.99									1	2
25	9.00-9.24										1
26	9.25-9.49										
27	9.50-9.74										1
28	9.75-9.99										
Total		55	64	50	82	56	70	63	64	81	64

*Readability scores determined by Spache and Dale-Chall formulas.

Note: Solid lines in columns indicate difficulty ranges of passages actually used in exercises. Not all passages shown here were used.

This is Table 3 of SSPED form 86 (page 11).

Table A-11. Number of Multiple-Choice Cloze Exercises by Difficulty Level of Passage, Domain, and Content Area

Difficulty level	Domain										Grand total
	Textual						Citizen		Consumer*	Reference *	
	Read./ Lit. *	Lang. Arts	Soc. Stud.*	Sci.	Math.	Total	Newspapers*	News Mags.*			
1	10		1			11					11
2	33		10			43					43
3	26	1	14			41					41
4	21	2	24			47					47
5	15	10	25	12	6	68				3	71
6	17	8	8	4	9	46				6	52
7	20	7	7		4	38				5	43
8	17	11	8	3	9	48				3	51
9	29	8	7	3	4	51	1			4	56
10	11	11	14	5	8	49	2			8	59
11	22	0	6	4	8	49	2		5	6	62
12	9	14	10	9	10	52	4		3	6	65
13	19	11	14	9	9	62	2	3	3	6	76
14	8	10	12	11	11	52	4	3	5	7	71
15	20	7	12		7	52	3	7	4	4	70
16	14	11	13	13	7	58	2	6	5	4	75
17	17	9	7	18	10	61	10	12	7	5	95
18	20	10	7	11	11	59	6	7	8	4	84
19	7	9	5	5	7	33	7	8	5	3	59
20	12	6	7	11	15	51	13	9	6	7	86
21	3	3	5	6	5	22	13	10	7	3	55
22	6	3	1	7	8	25	7	7	6	4	49
23			3	6	6	15	10	11	6	5	47
24				6	4	10	8	8	7	3	36
25				4		4	8	10	7	1	30
26				6	2	8	7	6	6		27
27				1		1	7	8	5		21
28							4	5	5		14
Total	356	160	220	160	160	1056	120	120	100	100	1496

This is Table 11 of the SSPED form 86 (page 33). The asterisks mark content areas we chose in our test construction. The two horizontal lines show the upper and lower difficulty levels we used in our selection process.

d) 13 was paired with 14.

e) 15, 16, and 17 were combined.

Any content category and difficulty level combination with fewer than four passages was eliminated because one selection was needed for each form of the test. Tables of random numbers were used to determine the content area (and difficulty level in combined strata) of the passages to be chosen.

Table A-12 shows the composition of each of the four forms. Newspapers and newsmagazine sources were not selected, because there were too few available passages relative to the other content areas. The passages were ordered by difficulty level in the test booklets. The first passage required 5 answers, the others 10 each, for a total of 75 items.

Passages were re-formatted so that the alternative answers were placed in a box at that point in the text. In the original exercises the alternative answers were placed at the bottom of the page. The original and revised forms are shown in Figures A-1 and A-2.

DEVELOPMENT OF THE MATHEMATICS TESTS

Domain Specification

The initial step in the specification phase for the mathematics tests was to determine the content corresponding to the abilities and courses of instruction of Title I students. This was done by: (1) asking the Policy Advisory Group to indicate boundary points of appropriate content for each domain; (2) reviewing existing collections of behavioral objectives and test items for the relevant grade levels; and (3) reviewing mathematics curriculum materials used in the public schools.

The definition of content for the practical mathematics domain was based primarily upon a review of curriculum materials used by the Job Corps (Bratten, Robertson, and Geddes, 1972). It was felt that this material would be particularly relevant to the population of Title I N or D students. The forms of the items created for the Job Corps were similar to those called "story problems" in other tests and curriculum materials. Table A-13 gives the final outline of the content elements and item formats for the mathematics domains.

Table A-12. Composition of Literal Comprehension Subtest, by Form

Content Area	Difficulty Level	Catalogue Number of Selected Passage			
		Form 1	Form 2	Form 3	Form 4
Social Studies	4	9	12	15	18
Reading/Literature	5	15	5	4	6
Social Studies	6	5	6	8	4
Reference	7	2	5	1	4
Reading/Literature	10	5	1	7	2
Consumer	11	3	4	5	1
Reading/Literature	14	1	7	5	2
Reading/Literature	17	10	15	3	11

They had tanks of air on their ___6___ . Rubber tubes came out of the ___7___ and went into the ___8___ masks to give them all the ___9___ they needed. These divers did not have to ___10___ to the surface of the ___11___ to get air.

Carlos ___12___ air. He did not ___13___ a tank to give him ___14___ under the water.

All at once Carlos ___15___ out.

(6)
1. umpires
2. experts
3. backs
4. ducks
5. grips

(7)
1. gloves
2. plows
3. beams
4. tanks
5. lakes

(8)
1. fat
2. misty
3. face
4. sunny
5. rare

(9)
1. air
2. shadow
3. ribbon
4. gold
5. loss

(10)
1. bicycle
2. go
3. pad
4. hop
5. salute

(11)
1. water
2. cornfield
3. ranger
4. dinner
5. bear

(12)
1. slung
2. nailed
3. smeared
4. needed
5. popped

(13)
1. have
2. spin
3. hear
4. trap
5. freeze

(14)
1. trust
2. bacon
3. ice
4. excitement
5. air

(15)
1. peeled
2. blacked
3. banged
4. stuck
5. camped

Figure A-1. Passage as Presented in Original Compilation

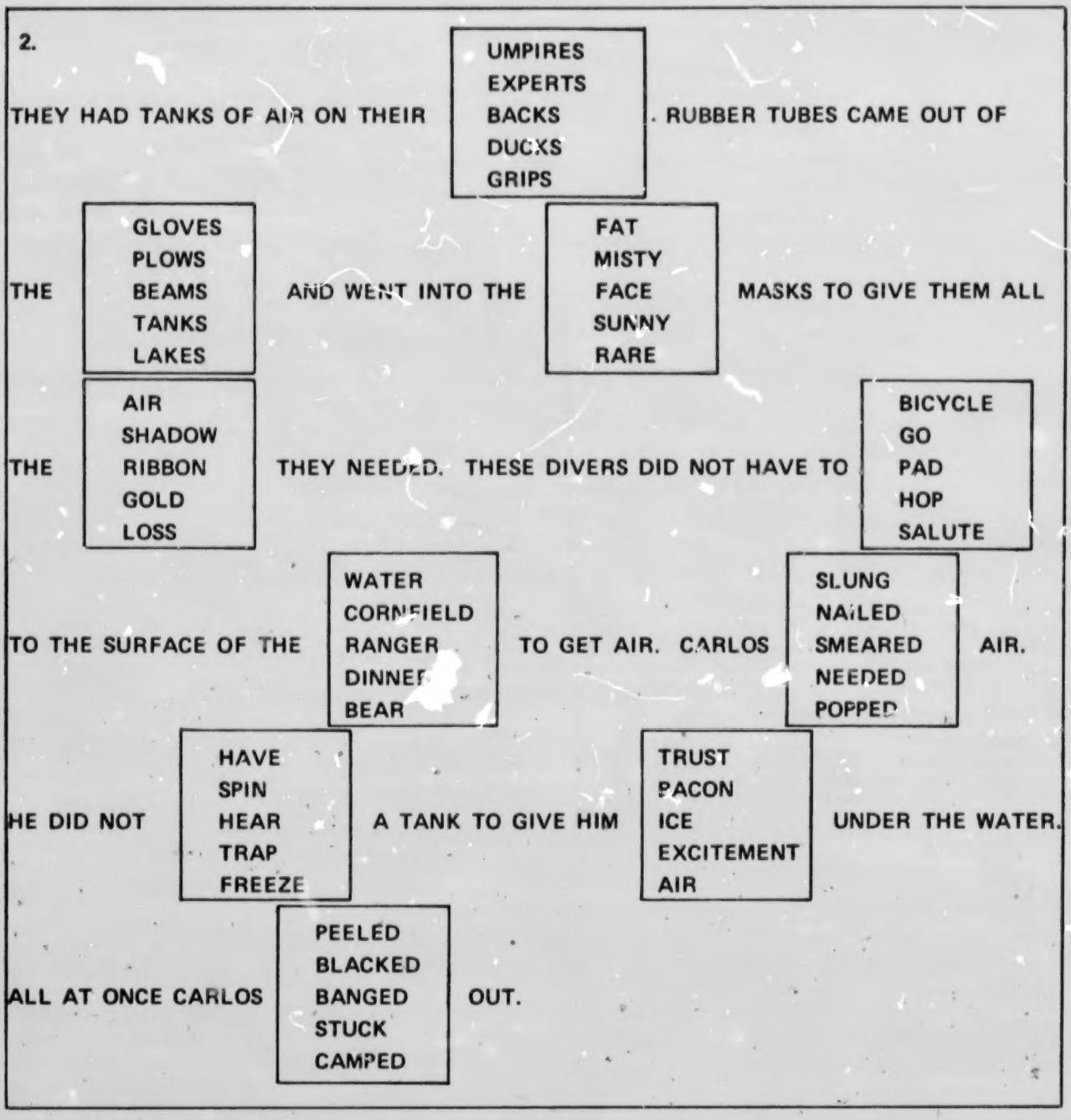

Figure A-2. Passage as Presented in DRT

Table A-13. Content Specification for Mathematics Domains

Domains	Content Elements of the Domain		Item Format
Addition	Facts:	Single digit addends	Column addition problems
	Integers:	Single digit to 3 digits; up to 4 addends	Column addition problems
	Decimals:	To hundredths place; mixed decimals; up to 2 addends	Column addition problems
	Fractions:	Common with like or unlike denominators; mixed fractions; up to 3 addends. Single digit denominators.	Row addition problems
	All numbers are zero or positive.		All problems require a constructed response.
Subtraction	Facts:	Single digit minuend and subtrahend	Column subtraction problems
	Integers:	Single digit to 3 digit minuends and subtrahends	Column subtraction problems
	Decimals:	To hundredths place; mixed decimals	Column subtraction problems
	Fractions:	Common with like or unlike denominators; mixed fractions. Single digit denominators.	Raw subtraction problems
	All results must be zero or positive.		All problems require a constructed response.
Multiplication	Facts:	Single digit multiplicands	Column multiplication problems
	Integers:	Single digit to 3 digits	Column multiplication problems
	Decimals:	To hundredths place; mixed decimals	Column multiplication problems
	Fractions:	Common or mixed fractions. Single digit denominators	Row multiplication problems
	All numbers are zero positive.		All problems require a constructed response.

Table A-13. Content Specifications for Mathematics Domains (Cont'd)

Domain	Content Elements of the Domain	Item Format
Division	Facts: Single digit divisors and quotients Integers: Single digit to 4 digit quotients Decimals: To hundredths place; mixed decimals Fractions: Common and mixed fractions. Single digit denominators All numbers are zero or positive.	Division problems using $a\overline{)b}$ Division problems using $a\overline{)b}$ Division problems using $a\overline{)b}$ Division problems using $b \div a$ All problems require a constructed response.
Practical Mathematics	Making Change Calculating Area Calculating Cost Comparing Prices Budget Calculating Interest Calculating Wages	Story problems for each content area All problems required a constructed response.

The constructed response format was chosen for two reasons. First, students would not be likely to obtain a correct response by guessing. Second, there would be no need to construct alternative responses.

To operate within the one-hour time constraint, 100 arithmetic problems and 12 story problems were given. Each mathematics subtest was given in a separate test booklet. Students were told that the nature of the problems was changing; the booklets were color-coded to emphasize these changes.

The structure of each test is given in Tables A-14 to A-18. Four versions of each item (one per form) were generated from tables of random numbers. It should be emphasized that the items in these tests were selected for their relevance to the N or D setting and may not be appropriate for other studies and evaluations.

Because the mathematics items required constructed responses, criteria had to be set for correct answers. For integer answers, commas were ignored (some students separated pairs of numbers rather than triples). If the result was a decimal number, the decimal had to be positioned correctly. A common fraction or mixed number did not have to be reduced (any common fraction or mixed number equal to the decimal equivalent of the correct answer was scored as correct). In division, remainders could be expressed in any of three notations:

- R notation, e.g., 15 ÷ 6 = 2 R 3
- Fraction notation, e.g., 15 ÷ 6 = 2 3/6 or 2 1/2
- Decimal notation, e.g., 15 ÷ 6 = 2.5

It was felt that the complexities introduced in scoring were more than compensated by the reduction in guessing and the elimination of complex rules for generating distractors.

Table A-14. Content Specification for Addition Subtest

Item Number(s)	Content
1 - 8	Add a number in the range (1-9) to another number in that range. No pair may be repeated. One pair must contain a zero.
9	Add a number in the range (1-9) to another number in that range. Answer must exceed 9.
10	Add three numbers in the range (1-9). Answer to exceed 10.
11	Add a number in the range (1-9) to a number the range (10-99). No regrouping required.
12	Same as item 11, but requires regrouping.
13	Add two numbers in the range (10-99). No regrouping of units required. No zeros in addends.
14	Like item 13, but requires regrouping in units and tenths position.
15	Add three numbers in the range (10-99). Regrouping required in both columns.
16	Add a number in the range (111-999); (no zeros in any position) to a number in the range (11-99; no zeros in either position). Regrouping of units not required. Regrouping of tens to hundreds required.
17	Like 16, but 3-digit number has a zero in tens place. No regrouping required.
18	Add two numbers in the range (111-999; no zeros in any position). Regrouping required in all columns.
19	Add a number in the range (111-999; no zeros) to a number in the range (11-99; no zeros) and to a number in the range (1-9) and to a number in the range (111-999; no zeros). Regrouping required in all columns.
20	Add 2 numbers in the range (.10 - .99). Correct answer is less than 1.00.

Table A-14. Content Specification for Addition Subtest (Cont'd)

Item Number(s)	Content
21	Like item 20, but answer must be between 1.00 and 1.99.
22	Add a number in the range (1.00 - 9.99) to a number in the range (.10 - .99) and to a number in the range (.01 - .09).
23	Add 2 fractions (each not reducible) with a common denominator in the range (3-9). Answer must be less than 1.
24	Like item 23, but answer must exceed 1.
25	Add three fractions with a common denominator in the range (4-9). Only one of the fractions may be reduced.
26	Add two fractions with different denominators in the range (3-9). One denominator is an integer multiple of the other. Answer is less than 1.
27	Add two fractions with different denominators in the range (3-9). Common denominator must be computed as the product of two denominators. Answer less than 1.
28	Add two mixed whole numbers and common fractions. Whole number parts in the range (1-9). Fraction parts have different denominators in the range (2-9). Common denominator must be computed as the product of the two denominators.

Table A-15. Content Specification for Subtraction Subtest

Item Number(s)	Content		
1 - 8	Subtract a number in the range (1-9) from a larger number in that range. One problem must have zero as the subtrahend. One problem must have zero as the answer.		
	Range of Minuend	Range of Subtrahend	Constraints
9	(10-99)	(1-9)	No borrowing
10	(10-99)	(1-9)	Borrowing
11	(11-99)	(10-98)	No borrowing
12	(11-99)	(10-98)	Borrowing
13	(111-999; no zeros)	(11-99; no zeros)	No borrowing
14	(111-999; no zeros)	(11-99; no zeros)	Borrowing in tens place only
15	(111-999; no zeros)	(11-99; no zeros)	Borrowing in units and tens places
16	(101-909; zero in hundreds place only)	(11-99; no zeros)	Borrowing required in units place
17	(111-999; no zeros)	(111-999; no zeros)	Borrowing in units and tens places
18	(.11-.99; no zeros)	(.11-.98; zeros)	No borrowing
19	(.11-.99; no zeros)	(.11-.98; zeros)	Borrowing; answer less than .10
20	(1.10-9.99; no zeros in tenths place)	(.10-.99)	Borrowing in tenths column
21	Fraction <1	Fraction <1	Common denominator of 3, 5 or 7
22	Fraction <1	Fraction <1	Denominator of one fraction is 2, 3 or 4; denominator of other is 4, 6, 8 or 9 - an integer multiple of the first

Table A-15. Content Specification for Subtraction Subtest (Cont'd)

Item Numbers(s)	Content		
	Range of Minuend	Range of Subtrahend	Constraints
23	Fraction <1	Fraction <1	Denominators in range (2-9). Common denominator must be computed as product of denominators.
24	Mixed Number < 10	Mixed Number < 10	Denominators in range (2-9) Common denominator must be computed as product of denominators. Minuend Fraction part > subtrahend fraction part.
25	Mixed Number < 10	Mixed Number < 10	Common denominator in range (2-9). Minuend fraction part < subtrahend fraction part.

All items 9-25 have answers greater than zero.

Table A-16. Content Specification for Multiplication Subtest

Item Number	Content		
1-8	Stratify single digit integers (2-9) into these groups: a: {2,3,4,5} b: {6,7,8,9} One item is zero times any digit in a or b. One item is one times any digit in a or b. Two items are products of single items from a. Two items are products of single items from b. Two items are products of one a item with one b item. No pairs may occur twice in one form.		
	Range of Multiplicand ①	Range of Multiplicand ②	Constraints
9	(10-99)	(2, 3, or 4)	No regrouping
10	(10-99)	(2-9)	Regrouping
11	(11-44; no digit greater than 4)	(11-44; no digit greater than 4)	No regrouping
12	(11-99; no zeros)	(11-99; no zeros)	-
13	(111-999; no zeros)	(11-99; no zeros)	-
14	(110-990; must have 0 in units column; no other zeros)	(11-99; no zeros)	-
15	(101-909; must have 0 in tens column; no other zeros)	(11-99; no zeros)	-
16	(111-999; no zeros)	(111-999; no zeros)	-
17	(111-999; no zeros)	(101-909; must have 0 in tens column; no other zeros)	-
18	(.2 - .9)	(.2 - .9)	Answer > .09
19	(.1 - .9)	(.1 - .9)	Answer > .10
20	(.10 - .99)	(.1 - .9)	-

Table A-16. Content Specification for Multiplication Subtest (Cont'd)

Item Number	Content		
	Range of Multiplicand ①	Range of Multiplicand ②	Constraints
21	(.11-9.99; no zeros)	(.11-.99; no zeros)	-
22	Fraction < 1	Fraction < 1	Different denominators in range (2-9). Can "cancel"
23	Fraction < 1	Fraction < 1	Different denominators in range (2-9). Can't "cancel"
24	Mixed number < 10	Mixed number < 10	Common denominator in range (2-9).

Table A-17. Content Specification for Division Subtest

Item Number		Content			
		Dividend	Divisor	Quotient	Constraints
		0	1-9	0	-
		(1-9)	(1-9)	1	-
1-8	Two items like this	(2-25)	(1-5)	(2-5)	No remainder
		(2-25)	(1-5)	(2-5)	No remainder
		(6-36)	(1-4)	(6-9)	No remainder
		(12-45)	(6-9)	(2-5)	No remainder
	Two items like this	(36-81)	(6-9)	(6-9)	No remainder
		(36-81)	(6-9)	(6-9)	No remainder
		No items may appear twice in one form.			
9		(10-99)	(2-9)	(2-9)	+ Remainder
10		(22-98; no zeros)	(2-9)	(11-49)	No remainder
11		(21-99; no zeros)	(2-9)	(10-49)	+ Remainder
12		(22-98; no zeros)	(11-49; no zeros)	(2-9)	No remainder
13		(23-99; no zeros)	(11-49; no zeros)	(2-8)	+ Remainder
14		(111-999; no zeros)	(11-99; no zeros)	(2-9)	No remainder
15		(111-999; no zeros)	(11-99; no zeros)	(11-99)	No remainder
16		(202-909; must have zero in hundreds place; no other zeros)	(2-9)	(101-404; must have zero in hundreds place; no other zeros)	No remainder
17		(102-408; must have zero in hundreds place; no other zeros)	(2-8)	(21-81)	No remainder Dividend forms single digit integer quotient with hundreds and tens places of minuend.

Table A-17. Content Specification for Division Subtest (Cont'd)

Item Number	Content			
	Dividend	Divisor	Quotient	Constraints
18	(111-999; no zeros)	(11-99; no zeros)	(2-9)	+ Remainder
19	(111-999; no zeros)	(11-99; no zeros)	(11-90)	+ Remainder
20	(.10-.81)	(2-9)	(.02-.09)	Quotient X 100 is integer
21	(.11-.99)	(.01-.09)	(10-99)	No remainder
22	(1.00-9.99)	(1.0-9.9)	(1.0-9.9)	Quotient X 10 is integer
23	Fraction < 1	Fraction < 1	-	Different denominators in the range 2-9. No chance to cancel
24	Mixed number < 10	Mixed number < 10	-	Different denominators in the range 2-9.

Table A-18. Content Specification for Practical Mathematics Subtest

Item Number	Content
1	Making change: Converting different coins into one type of coin
2	Making change: Amount of money to be returned from a purchase. Amount offered is $5, $10, $20.
3	Calculating Cost: Given a unit cost state how much several units will cost.
4	Comparing Prices: Given a unit cost for one item and the cost of several units of another item state which is the best buy.
5	Measuring Area: Area of wall to be painted.
6	Measuring Area: Given length and width of field and coverage of one bag of seed, state number of bags of seed needed to cover field.
7	Calculating Cost: Given cost of n items compute the cost of m items.
8	Budget: Given monthly income and percentage distribution in 3 categories, state dollar amount paid for one of the categories.
9	Budget: Given monthly pay and percent of monthly pay spent on rent and food, state the dollar amount remaining each month.
10	Interest: Given the dollar amount of a loan and the interest rate, state the amount of interest due at the end of one year.
11	Interest: Given the dollar amount of a bank deposit and the yearly interest rate, state the total amount of money availalbe in the account after one year.
12	Wages: Given a hourly pay rate, the hours worked each week and the weekly tax withholding, state the amount of weekly take home pay.

The numbers used in these items varied across forms. The numbers were chosen to provide for easy computation.

APPENDIX B

HIERARCHICAL STRUCTURE OF THE DOMAIN-REFERENCED TESTS

In Chapter III the hierarchical nature of the tests was explored by examining the means of the operation and content types. The means clearly confirmed the order of difficulty. Although the means were suggestive of a hierarchical structure, the structure was not absolutely demonstrated because the tables did not show whether students who mastered later subtests, i.e., obtained perfect scores, answered all earlier items correctly. Table B-1 shows the proportion of correct responses on preceding subtests for mastery of each of the later subtests. In general, students who achieved perfect scores on later subtests had mastered preceding subtests. The fact that decimals do not form a complete prerequisite to fractions is born out by the less than perfect scores on decimals for students who had mastered fractions perfectly. The prerequisite of facts for integers, and facts and integers for decimals and fractions, is clearly shown. The results in Table B-1 then, further demonstrate the hierarchical structure of the arithmetic tests.

The pooled within-institution correlations among subtests also help to confirm the hierarchical ordering of the subtests within a test. The intercorrelations are presented in Table B-2 through B-5. If students do acquire new skills to perform more complex subtests, then the correlations between adjacent subtests should be higher than correlations between subtests separated in the ordering. That is, correlations should grow progressively smaller as one moves away from the main diagonal (simplex pattern). This pattern is demonstrated almost perfectly for all tests. Across all the correlations in the four tests there are only two correlations that do not conform to this pattern--one in addition and one in multiplication. These patterns of correlations constitute additional evidence to verify the hierarchical order among subtests.

The pooled within-institution correlations were also computed among tests to explore more fully whether the skills of addition, subtraction, multiplication, and division were hierarchically ordered. As shown in Table B-6, scores on adjacent tests were more highly correlated than tests separated in the ordering.

Analyses of the language arts domain-referenced tests were conducted using the same procedures as described above. The intercorrelations among the passages in Table B-7 display a hierarchical order corresponding to the order of presumed difficulty of the passages.

Table B-1. Mean Proportion Correct on Arithmetic Subtests for Students Who Achieved Perfect Scores on Later Subtests

Subtest	Proportion Correct			
	Addition	Subtraction	Multiplication	Division
Perfect scores on Integers				
Facts	.99	.98	.98	.99
Perfect scores on Decimals				
Facts	.99	.99	.98	.97
Integers	.96	.94	.84	.73
Perfect scores on Fractions				
Facts	.99	1.00	.99	.98
Integers	.97	.97	.85	.74
Decimals	.87	.92	.66	.69

Table B-2. Intercorrelations Among Addition Subtests at Wave 1*

	Facts	Integers	Decimals	Fractions
Facts	1.00			
Integers	.17	1.00		
Decimals	.08	.33	1.00	
Fractions	.09	.22	.37	1.00

Table B-3. Intercorrelations Among Subtraction Subtests at Wave 1*

	Facts	Integers	Decimals	Fractions
Facts	1.00			
Integers	.28	1.00		
Decimals	.19	.46	1.00	
Fractions	.13	.36	.43	1.00

*Only students from the 30 institutions selected at random were included in this analysis. The results for students in all 40 institutions would be much the same.

Table B-4. Intercorrelations Among Multiplication Subtests at Wave 1*

	Facts	Integers	Decimals	Fractions
Facts	1.00			
Integers	.54	1.00		
Decimals	.30	.50	1.00	
Fractions	.30	.43	.32	1.00

Table B-5. Intercorrelations Among Division Subtests at Wave 1*

	Facts	Integers	Decimals	Fractions
Facts	1.00			
Integers	.49	1.00		
Decimals	.33	.74	1.00	
Fractions	.18	.39	.52	1.00

*Only students from the 30 institutions selected at random were included in this analysis. The results for students in all 40 institutions would be much the same.

Table B-6. Intercorrelations Among Arithmetic Tests at Wave 1*

	Addition	Subtraction	Multiplication	Division
Addition	1.00			
Subtraction	.76	1.00		
Multiplication	.62	.73	1.00	
Division	.60	.71	.76	1.00

Table B-7. Intercorrelations Among the Eight Literal Comprehension Passages at Wave 1*

Passage	1	2	3	4	5	6	7	8
1	1.00							
2	.69	1.00						
3	.69	.86	1.00					
4	.62	.78	.85	1.00				
5	.60	.75	.81	.85	1.00			
6	.50	.69	.78	.78	.87	1.00		
7	.49	.67	.74	.74	.83	.87	1.00	
8	.40	.59	.63	.61	.71	.75	.79	1.00

*Only students from the 30 institutions selected at random were included in this analysis. The results for students in all 40 institutions would be much the same.

APPENDIX C

RESULTS OF THE FACTOR ANALYSIS OF STUDENT ATTITUDE SCALES

All items measuring self-attitudes, attitudes toward school and learning, and perceptions of the environment were combined on the Student Attitude Measures I and II instruments. Because of the number of dimensions included in this instrument--and anticipated problems in maintaining student attention--attitudinal items were administered at two testings on consecutive days. To reduce the likelihood of response sets, scale items were intermixed and split between forms.

All hypothesized scales were subjected to a factor analysis in an attempt to investigate the dimensionality and factorial structure presumed to be measured by conceptual areas. Response data from the 30 randomly selected sites were factor analyzed for the first test administration. This group of students represented either the entire educational population or a random sample of the educational population at each institution (N = 886). Recently institutionalized students at the third site visit who had not been given attitudinal items during the first test administration form a cohort comparable to the initial sample (N = 406). Results of the factor analysis from the initial group were compared to results from the group at the third testing. The latter factor analysis confirmed the structure revealed by the initial factor analysis. Aside from the slightly lower factor loadings and reliability estimates for the third testing group, all factor scales and item loadings were similar for the two analyses. Factor analytic results reported in this section reflect an analysis based on the combined test 1 and test 3 group data. This combined group represents each student for whom the Student Attitude Measures were administered on at least one occasion.

Factors were extracted with a principal factors solution. Item communalities were initially estimated by the multiple R^2 and then iterated to a thirteen-factor solution. The factor-extraction cut-off resulting in the retention of these factors was established at 1.0. Although the initial orthogonal rotation had yielded reasonably clear factor structures, an attempt was made to obtain

clearer and more reliable independent factor scales. Items with low communalities (<.20) were eliminated prior to this rotation, as were items with low factor loadings (<.25), and items loading significantly on more than one factor.

The results of this final solution are presented in Table C-1.* Self-attitude dimensions obtained were self-esteem, locus of control, sociability, and expectations for post-release success or stigmatization. Aside from several self-esteem items that loaded significantly on more than one factor, the structure of these dimensions was confirmed to be as hypothesized.

The dimensions identified by the factor analysis of attitudes toward school and task-related activities, such as reading and mathematics, differed somewhat from hypothesized dimensions. The dimension attitudes toward school split into two dimensions: school involvement and liking school. Attitudes toward reading tasks and math tasks form two separate dimensions, as anticipated.

The final area for attitude assessment that had been conceptualized was the psychosocial perceptions of institutional environments. Hypothesized dimensions included expressiveness, affiliation, rule clarity, practical orientation, order and organization, teacher support, and the importance of education to institutional staff. Prior use of these dimensions had failed to indicate which conceptual areas were analytically distinct. Results of the factor analysis of these perceptual dimensions of environment failed to substantiate the system maintenance dimensions of order, organization, and rule clarity (Moos, 1972).

The second basic dimension investigated by means of psychosocial perceptions was the personal development dimension. Structures emerging in this area were 'practical orientation of learning', and 'school involvement'. Hypothesized concepts such as 'personal problem orientation' and 'expressiveness' failed to emerge.

---

*Items were scored "1" for responses indicating positive attitudes and "0" for responses indicating negative attitudes.

Table C-1. Factor Scales and Loadings

	SELF-ESTEEM	LOCUS OF CONTROL	SOCIABILITY	PEER AFFILIATION	SCHOOL INVOLVEMENT	LIKING OF SCHOOL	ATTITUDES/ MATH	ATTITUDES/ READING	TEACHER SUPPORT	PRACTICAL ORIENTATION OF LEARNING	RELEASE EXPECTATIONS/ STIGMATIZATION
SELF-ESTEEM:											
There are a lot of good things about myself.	.30	.02	.14	.06	.12	-.00	.04	.10	.03	.00	.06
Most of the time I'm proud of what I do.	.54	.02	.06	.12	.06	-.03	.04	.06	.06	.04	.04
I usually know what I want to do.	.32	-.01	.00	.13	.06	.02	.12	-.02	.06	.06	-.02
I usually feel good about myself.	.00	.05	.06	.00	.06	-.01	.04	.04	.07	.04	.02
I'm not very proud of myself.	-.48	.27	.01	-.03	.00	.04	.03	-.01	-.01	-.01	.03
I am happy with myself.	.00	.06	.11	.06		.01	.02	-.01	.10	.01	.11
I can do most things as well as other people.	.35	.10	.10	.12	-.04	-.03	.13	.06	.06	-.01	-.01
LOCUS OF CONTROL:											
I only work when my teachers ask me to.	.05	.35	-.04	.12	.14	.14	.10	.06	-.00	.10	.06
It's almost useless to try to learn in school because I can't do the work.	.12	.00	-.03	.05	.06	.11	.10	.00	-.02	.03	-.02
Being sent here has ruined my whole life.	.14	.01	.03	.00	-.04	.06	.00	-.01	.06	.03	.20
Many times I feel that I have little say over things that happen to me.	.14	.36	.07	.01	.00	.15	.03	.05	.01	-.03	.08
When bad things are going to happen, they just happen no matter what I try to do to stop them.	.16	.45	.01	-.04	-.04	.15	.12	.00	.03	.00	.14
Teachers here order me around a lot in class.	-.04	.00	-.07	.24	.14	.13	.03	-.03	.12	.04	.04
When I do something wrong, there is very little I can do to make it right.	.05	.58	.05	.01	-.03	.06	-.05	.12	.05	.02	.06
One of the best ways to handle most problems is just not to think about them.	-.05	.04	-.01	-.02	.00	.1	-.03	.12	-.01	-.03	-.17
When people are mean to me, it is usually for no reason at all.	.06	.42	.04	.06	-.05	-.07	-.03	-.03	.06	-.01	.04
No matter how hard I try, no one notices my efforts.	.00	.53	.10	.01	-.03	.00	.06	-.02	.10	.03	-.07
I am often blamed for things that aren't my fault.	.07	.43	.00	.04	.06	.13	.04	-.01	.06	.02	.04
It usually doesn't pay to try hard because things never turn out right anyway.	.12	.38	.07	.06	.01	.00	.04	-.01	.05	.00	.02
SOCIABILITY:											
I have a lot of friends.	.03	.01	.53	-.04	-.02	-.04	.05	.10	.22	.06	.00
I'm easy to like.	.09	.10	.53	.14	.06	.00	.06	.09	-.01	.02	.11
Most people like to be around me.	.01	.07	.00	.11	.06	-.01	.05	.07	.04	.06	.01
I'm a lot of fun to be with.	.27	.03	.47	.00	.04	.03	.02	.07	.04	-.02	.00
PEER AFFILIATION:											
Students here get along well with other students.	.13	.00	.10	.48	.01	.06	-.01	.02	.12	.03	-.01
I get along with other students in my classes.	.20	.12	.22	.42	-.02	.07	-.05	.10	.07	.06	-.01
Students of different races get along with each other here.	.06	.05	.08	.40	-.06	.01	-.02	-.01	.05	-.0	.13
SCHOOL INVOLVEMENT:											
I work hard in class.	.00	.05	.04	-.05	.52	.17	.00	.04	.00	.00	-.03
I try hard to do better in school.	.06	-.04	.00	-.01	.02	.10	.15	.10	.07	.25	.01
I'm doing the best work I can in school.	.07	-.15	.02	.02	.57	.06	.01	.02	.00	.10	.07
LIKING OF SCHOOL:											
I would like to stop taking some of my classes.	.07	.10	-.05	.12	.17	.34	.20	.11	.12	.10	-.05
My classes are usually dull and boring.	-.03	.19	-.01	.07	.19	.45	.06	.13	.23	.16	.00
I am usually unhappy when I am in school.	.05	.21	.05	.07	.12	.54	.14	.06	.07	.12	.04
Most days I do not want to go to school.	.03	.24	-.02	.03	.21	.47	.24	.08	.05	.06	.07
I often get upset in school	.00	.29	.00	.07	-.02	.41	.07	.00	.05	.01	-.02
ATTITUDES TOWARD MATH:											
I don't like multiplying numbers.	.07	.10	.04	-.03	-.02	.11	-.58	-.01	.03	.01	.01
I like doing problems with fractions.	.07	.04	.00	-.03	.12	.11	.50	.00	.11	.02	.00
I like dividing numbers.	.10	.00	.03	.02	.00	.10	.00	.13	.04	.05	.05
ATTITUDES TOWARD READING:											
I like reading stories.	.05	.02	-.04	.01	.03	.10	.07	.57	.01	.03	.01
I like studying the meaning of words.	.11	.04	-.05	.00	.10	.14	.27	.43	.04	.06	.04
I like reading newspapers and magazines.	.01	.06	.06	.0	.00	.03	.00	.47	.06	.01	.01
TEACHER SUPPORT:											
Most of my teachers here are fair.	.05	.10	-.04	.10	.06	.10	.06	-.03	.30	.10	-.05
Teachers here go out of their way to help me.	-.01	.05	.03	.02	.12	.05	.03	-.01	.60	.12	.06
Teachers here care what happens to me when I leave here.	-.05	.11	.05	.06	.02	.04	.06	-.03	.60	.13	.11
If my teacher does something I don't like, I can tell him/her about it.	.06	.07	.01	.14	.04	.02	.13	.02	.36	.03	-.00
I can talk to teachers about things that bother me.	.06	-.03	.01	.00	.04	.13	.04	.11	.56	.10	.35
My teachers tell me when I am doing well.	.12	.04	.04	.06	.04	.00	.02	.03	.46	.13	.03
Most people like to help me out when I need help.	.14	.06	.06	.10	.02	-.01	.01	.00	.43	.00	.13
PRACTICAL ORIENTATION OF LEARNING:											
I am learning a lot in my classes.	.04	.06	.01	.10	.02	.10	.02	.02	.32	.44	.03
I am learning things I'll need to know when I leave here.	.02	.04	.03	.02	.12	.05	.06	.02	.17	.73	.06
What I have learned in school will help me when I leave here.	.05	.07	-.01	-.02	.12	.00	.06	.03	.22	.73	.04
I have learned things in school here that will help me when I leave here.	.00	.04	-.01	.02	.15	.11	.01	.06	.32	.61	.06
EXPECTATIONS FOR RELEASE SUCCESS AND STIGMATIZATION:											
It will be hard for me to stay out of trouble with the law now that I have been in a place like this.	.06	.30	-.04	.14	.05	.07	.08	.06	-.06	.04	-.35
I think people will give me a fair chance when I leave here as long as I stay out of trouble.	.05	.02	.12	.06	.04	.04	.03	.07	.16	.12	.36
I think that having been here will hurt my chances for getting a good job when I leave here.	.08	.35	.10	-.06	.03	.05	-.01	-.03	.10	.03	-.37
Most people on the outside don't give someone who has been in a place like this a fair chance.	-.02	.32	.06	-.07	-.01	.03	-.02	-.01	.13	.04	-.36

Relationship dimensions were the final aspect of perceived environments investigated. These dimensions included relationship aspects such as involvement in the environment, interpersonal support, and staff support. Factors identified that were indicative of the relationship dimension include school involvement, staff support, and peer affiliation. Expressiveness, a dimension hypothesized by Moos to be reflective of the personal development dimension, loaded on the teacher support dimension. Although an attempt was made to measure the importance of education in the institutional environment, these items failed to form a scale. Results on the factor analysis are reported in Table C-1.

One check upon the reasonableness of the factor-analytic solution is to obtain scores for each of the independent scales and to intercorrelate the scales. Scale intercorrelations should be sufficiently small so that there is reason to believe the scales provide assessments of different student attitudes. Table C-2 provides the scale intercorrelations as well as reliability estimates for each of the scales. For most scales, estimated reliabilities were acceptable. The scale 'release expectations and stigmatization' demonstrated relatively low reliability and correlated moderately with the locus-of-control scale. This argues for cautious use of the scale. Locus-of-control also demonstrated a moderate relationship to 'liking of school'. It should be noted that a possible reason why these scales demonstrate less independence is the fact that the scales are comprised of items with negative content. Students may be responding more consistently to negative items regardless of the construct measured.

Table C-2. Factor Scale Intercorrelations (below diagonal) and Kuder-Richardson Reliability Estimates (diagonal)

	Self-Esteem	Locus of Control	Sociability	Peer Affiliation	School Involvement	Liking of School	Attitudes/ Math	Attitudes/ Reading	Teacher Support	Practical Orientation of Learning	Release Expectations/ Stigmatization
Self-Esteem	.71										
Locus of Control	.27	.78									
Sociability	.30	.16	.65								
Peer Affiliation	.25	.22	.22	.56							
School Involvement	.16	.02	.10	.04	.64						
Liking of School	.18	.41	.10	.17	.28	.69					
Attitudes Toward Math	.19	.19	.10	.05	.20	.33	.69				
Attitudes Toward Reading	.17	.16	.05	.07	.19	.25	.27	.52			
Teacher Support	.21	.24	.21	.25	.19	.31	.17	.13	.72		
Practical Orientation of Learning	.17	.18	.12	.12	.38	.36	.17	.17	.46	.80	
Release Expectations and Stigmatization	.19	.46	.18	.14	.05	.25	.11	.08	.24	.19	.48

APPENDIX D

AVERAGE STUDENT BACKGROUND CHARACTERISTICS AND ATTITUDES FOR EACH OF 30 RANDOMLY SAMPLED INSTITUTIONS

Table D-1. Distribution of Student Characteristics Within Institutions
(Thirty Randomly Sampled Institutions)

Institution	n	Age		Sex		Grade		Number of Schools Attended		Months left in Institution		Number of times Institutionalized	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	71	15.3	1.3	1.3	.5	8.8	1.6	6.5	2.5	3.7	3.2	3.7	3.3
2	51	14.5	1.9	1.6	.5	7.9	2.0	6.1	2.7	10.7	11.8	3.3	2.5
3	19	16.0	.8	1.8	.4	8.8	1.6	5.6	2.6	13.5	24.5	1.4	1.3
4	126	15.4	1.5	2.0	.1	9.2	1.3	5.2	2.0	6.0	7.3	2.7	2.4
5	76	16.1	1.9	1.8	.4	9.1	1.4	5.4	2.2	4.7	4.5	3.0	2.7
6	82	15.3	1.0	1.9	.3	8.4	1.2	4.3	2.3	6.8	7.2	2.9	2.9
7	68	14.9	1.0	1.8	.4	8.6	1.5	5.5	2.3	5.1	5.0	3.3	2.7
8	63	15.1	1.2	1.8	.4	8.5	1.4	4.7	2.3	3.0	2.7	2.1	1.9
9	80	18.3	1.4	2.0	.1	9.8	1.4	4.1	2.6	5.5	6.8	1.4	1.3
10	136	14.8	1.0	1.7	.5	8.7	1.2	4.7	2.6	4.4	6.2	1.9	1.6
11	97	16.0	1.0	1.8	.4	9.2	1.3	5.2	2.7	3.8	5.3	2.2	1.9
12	42	16.4	1.1	2.0	.0	9.8	1.0	6.1	2.5	2.9	2.1	2.5	2.9
13	69	14.9	1.4	1.7	.5	8.5	1.3	3.7	2.6	1.6	2.7	1.6	1.3
14	79	13.2	1.0	2.0	.0	7.1	1.4	4.2	2.6	4.9	4.5	2.1	2.2
15	56	16.8	.7	2.0	.0	10.0	1.0	5.2	2.3	4.4	4.3	2.1	1.7
16	36	16.9	.7	2.0	.0	9.9	1.1	5.3	2.9	4.0	2.4	2.2	1.5
17	76	16.7	1.2	1.9	.3	9.7	1.3	6.2	3.0	3.8	3.0	2.5	2.3
18	64	16.6	2.0	1.6	.5	9.2	1.8	5.3	2.6	7.6	14.6	2.8	3.0
19	104	18.7	1.6	2.0	.0	10.5	1.5	5.7	2.6	8.6	7.3	1.9	2.0
20	97	15.0	1.3	1.0	.0	9.1	1.3	6.1	2.5	3.5	2.7	1.8	1.4
21	85	19.2	1.1	2.0	.2	9.6	1.5	4.7	2.3	11.6	9.4	1.2	1.4
22	90	22.7	4.5	2.0	.1	9.7	1.4	5.0	2.4	20.6	24.2	2.0	1.9
23	40	21.8	2.9	2.0	.0	9.2	1.9	5.2	2.7	26.6	24.9	2.3	2.2
24	95	21.6	3.7	2.0	.0	9.1	2.0	3.9	2.2	23.1	26.0	2.3	2.4
25	61	20.1	2.0	2.0	.0	9.2	1.3	4.5	2.2	5.1	6.7	2.5	2.2
26	45	22.2	2.0	2.0	.0	8.7	2.1	4.8	2.6	20.5	22.1	2.1	1.6
27	122	20.1	1.9	2.0	.1	9.6	1.7	5.1	2.3	12.0	10.0	2.0	2.1
28	92	23.9	5.0	2.0	.1	10.0	1.3	4.5	2.2	16.2	15.3	2.0	2.3
29	63	20.5	2.0	2.0	.1	9.3	1.6	5.0	2.2	25.4	34.6	2.3	2.8
30	64	21.3	2.6	2.0	.1	10.1	1.7	4.0	2.4	5.3	4.5	1.3	1.4

Table D-2. Percent of Students in Each Category of Reason for Institutionalization at Each Institution (Thirty Randomly Sampled Institutions)

Institution	Reason for Institutionalization: Neglected or Dependent	Truancy, Runaway, Incorrigible	Burglary, larceny, auto theft, Arson, Fraud, Embezzlement	Assault, Battery, Sexual Assault Homicide, Manslaughter, Rape	Drug Offenses
1	2	74	7	7	8
2	63	28	0	4	4
3	7	14	43	29	7
4	5	8	65	15	7
5	5	20	53	16	5
6	0	15	58	24	3
7	3	10	43	38	8
8	4	21	57	11	7
9	1	3	56	24	16
10	4	31	45	14	6
11	1	14	55	22	8
12	3	7	55	21	14
13	2	48	45	5	2
14	3	16	65	12	4
15	5	2	74	17	2
16	0	5	58	16	21
17	0	25	53	18	4
18	0	23	36	30	11
19	2	2	32	58	5
20	4	80	2	4	9
21	0	0	56	40	4
22	0	0	24	76	0
23	3	0	13	82	0
24	0	2	24	62	11
25	2	4	66	25	4
26	2	0	15	66	17
27	1	1	56	33	9
28	0	0	33	58	10
29	0	2	14	71	12
30	2	0	61	26	11

Table D-3. Percent of Students in Each Ethnic Category at Each Institution (Thirty Randomly Sampled Institutions)

	Ethnic Background				
Institution	American Indian or Alaskan Native	Asian or Pacific Islander	Black	White	Hispanic
1	11	0	13	75	1
2	14	0	6	70	10
3	0	0	32	26	42
4	9	0	5	80	6
5	12	1	8	72	6
6	5	1	44	38	13
7	13	0	32	41	14
8	5	2	29	46	19
9	14	0	15	71	0
10	7	0	56	38	0
11	10	0	43	44	4
12	6	0	36	58	0
13	1	1	67	28	3
14	3	0	68	29	0
15	7	2	36	54	2
16	8	0	8	69	14
17	11	2	16	46	25
18	16	0	11	38	35
19	11	0	52	17	21
20	14	0	2	72	11
21	6	0	46	36	12
22	9	0	49	40	2
23	5	0	73	22	0
24	3	0	72	24	1
25	7	0	34	59	0
26	0	0	91	9	0
27	8	1	55	35	1
28	2	3	40	52	2
29	4	0	52	40	4
30	3	0	20	77	0

APPENDIX E

DISTRIBUTION OF RESPONSES OF TITLE I AND REGULAR EDUCATION TEACHERS

Table E-1. Summary of the Mean Responses of Teachers of Basic Skills at 30 Randomly Selected Institituions

Variable	n	Mean	SD
Percentage of time devoted to classroom responsibilities	29	85.8	11.2
Size of group tasks usually assigned to. (100 = whole class, 200 = small groups, 300 = individual students).	29	230.4	35.8
Number of frustrations concerning teaching	29	1.9	.8
Number of years teaching experience in correctional institutions.	29	.7	1.5
Number of years teaching experience in other settings.	29	4.8	3.8
Hours of staff development/inservice training during the past year.	29	27.4	27.7
Degree to which teachers make decisions (3 = teachers, 1 = others)	29	2.7	.7
Degree to which education personnel make decisions (4 = education, 1 = others)	29	3.0	.7
Coursework in remedial instruction (0 = no, 1 = yes)	29	.6	.3
Coursework in diagnosis of special learning problems. (0 = no, 1 = yes)	29	.6	.3
Frequency of updating individual student performance objectives. (1 = once a year, 5 = at each class meeting)	29	3.7	.9
Degree of individualization (1 = low, 2 = high)	29	1.7	.1
Extent to which students participate in their own discipline (1 = never, 5 = more than 75% of the time)	29	3.5	.8
Frequency of staff meetings (1 = once a year, 4 = weekly)	29	3.3	.6
Degree EPA encourages discussions of problems (1 = actively encourages, 4 = discourages)	29	1.7	.6

Table E-1. Summary of the Mean Responses of Teachers of Basic Skills at 30 Randomly Selected Institutions (Cont'd)

Variable	n	Mean	SD
Frequency of feedback from EPA (1 = almost never, 4 = monthly)	29	2.9	.7
Helpfulness of EPA feedback (1 - very unhelpful, 4 = very helpful)	29	3.1	.4
Possession of state teaching credential or license (1 = yes 2 = no)	29	1.1	.2
Cooperation between students and teachers 1 = very poor, 4 = very good)	29	3.5	.3
Cooperation between students and treatment staff (1 = very poor, 4 = very good)	29	3.1	.4
Cooperation between students and cottage staff (1 = very poor, 4 = very good)	25	3.0	.5
Cooperation between students and custody staff (1 = very poor, 4 = very good)	28	2.8	.5
Cooperation between students and other residents (1 = very poor, 4 = very good)	29	3.1	.4
Students like academic education classes (1 = strongly disagree, 4 = strongly agree)	29	2.8	.3
Students put a lot of energy into their classes (1 = strongly disagree, 4 = strongly agree)	29	2.4	.5
Students are careful about what they say around teachers (1 = strongly disagree, 4 = strongly agree)	29	2.4	.5
Students have a say about what they do in classes (1 = strongly disagree, 4 = strongly agree)	29	2.7	.4
Classes are disorganized (1 = strongly disagree, 4 = strongly agree)	29	1.8	.4
Most students can potentially achieve at their grade level (1 = strongly disagree, 4 = strongly agree)	29	2.7	.4
Teachers will tell students why their program is changed (1 = strongly disagree, 4 = strongly agree)	29	3.1	.4
Students never know when they will be transferred (1 = strongly disagree, 4 = strongly agree)	29	2.2	.6

Table E-1. Summary of the Mean Responses of Teachers of Basic Skills at 30 Randomly Selected Institutions (Cont'd)

Variable	n	Mean	SD
Support given by Title I Teachers (1 = never supportive, 4 = always supportive	29	3.7	.3
Support given by EPA (1 = never supportive, 4 = always supportive)	29	3.5	.4
Support given by facility administrators (1 = never supportive, 4 = always supportive)	29	3.3	.5
Support given by custody staff (1 = never supportive, 4 = always supportive)	28	3.2	.5
Support given by cottage staff (1 = never supportive, 4 = always supportive)	24	3.2	.5
Involvement of facility director (1 = not involved, 4 = always involved)	29	2.5	.5
Involvement of EPA (1 = not involved 4 = always involved)	29	3.5	.4
Agreement with decisions for entire facility (1 = never agree, 4 = always agree)	29	2.9	.2
Agreement with decisions for entire education (1 = never agree, 4 = always agree)	29	2.9	.2
Frequency of use of teacher-made tests (1 = never, 5 = at each class session)	28	3.5	.6
Frequency of use of standardized tests (1 = never, 5 = at each class session)	29	2.1	.6
Frequency of use of special diagnostic tests (1 = never, 5 = at each class session)	27	2.3	.6
Frequency of use of programmed instruction worksheets (1 = never, 5 = at each class session)	28	3.6	.8
Frequency of use of measures of student attitudes/ behavior (1 = never, 5 = at each class session)	28	3.7	.9

The data in this table summarize the distribution of means of the responses obtained from teachers at each of the randomly selected institutions.

No teacher data were obtained from one of the institutions.

Table E-2. Summary of the Mean Responses of Title I Reading Teachers at 30 Randomly Selected Instititutions

Variable	n	Mean	SD
Coursework or training in behavior modification procedures, building student's self-esteem, diagnostic-prescriptive approaches, use of interest/learning centers, individualized approaches, intergroup relations, objectives-based approaches, use of multiethnic material, recording student progress, self-instructional techniques, use of student performance contracts, testing and evaluation of students, use of aides and volunteers, survey of new methods and products, student tutoring (o = none, 15 = all)	28	7.8	4.1
Use of approach that includes reading in content fields (1 = major approach, 4 = never used)	28	2.1	.7
Hours per week of assistance from special reading teacher	28	3.1	7.7
Hours per week of assistance from teaching aide	29	10.3	14.5
Hours per week of assistance from student aide.	28	3.2	7.7
Hours per week of assistance from student teacher	28	.9	3.2
Frequency of telling students their strengths and weaknesses (1 = never, 6 = at each class meeting)	28	5.4	.6
Frequency of grading students (1 = never, 6 = at each class meeting)	28	3.5	1.9
Use of phonics approach (1 = major approach, 4 = never used)	28	2.0	.8
Use of literal comprehension (1 = major approach, 4 = never used)	28	1.7	.6
Use of modified alphabet system (1 = major approach, 4 = never used)	28	2.8	.5
EPA thinks basic skills should be set aside until students are ready to learn (1 = yes, 2 = no)	28	1.8	.4
Frequency of assistance from speech therapist (1 = never, 6 = every class)	26	1.3	.7

Table E-2. Summary of the Mean Responses of Title I Reading Teachers at 30 Randomly Selected Institutions (Cont'd)

Variable	n	Mean	SD
Frequency of assistance from media specialist (1 = never, 4 = every class)	27	1.7	1.1
Frequency of use of special equipment (4 = never, 24 = every class)	28	14.2	6.8
Frequency of use of instructional materials (7 = never, 42 = every class)	26	33.0	5.7
Amount of advance planning (1 = low, 3 = high)	28	2.0	.8
Involvement of EPA (1 = none, 4 = much)	28	3.0	.7
Frequency of assistance from counselor, psychologist (1 = never, 4 = weekly)	27	2.3	1.2

The data in this table summarize the distribution of means of the responses obtained from teachers at each of the randomly selected institutions.

No teacher data were obtained from one of the institutions.

One additional institution had no Title I reading teacher.

Table E-3. Summary of the Average Responses of Title I Mathematics Teachers at 30 Randomly Selected Institutions

Variable	n	Mean	SD
Time devoted to teaching per week	25	18.8	9.7
Hours per week assistance from special math teachers	17	4.1	10.0
Hours per week assistance from teacher aide	23	13.3	15.2
Hours per week assistance from volunteer adult aide	19	1.0	2.2
Hours per week assistance from student aide	21	2.3	3.9
Frequency of assistance from counselor, psychologist (1 = never, 4 = weekly)	23	1.9	1.1
Frequency of use of special equipment (5 = never, 30 = every class)	24	13.4	7.1
Frequency of use of instructional materials (5 = never, 30 = every class)	22	21.7	3.8
Amount of advance planning (1 = low, 3 = high)	24	2.1	.8
Frequency of grading (; = never, 6 = every class)	23	4.1	2.0

The data in this table summarize the distribution of means of the responses obtained from teachers at each or the randomly selected institutions.

No teacher data were obtained from one of the institutions.

Four institutions had no Title I mathematics teacher.

APPENDIX F

DISTRIBUTION OF TEACHER RESPONSES WITHIN INSTITUTIONS

Table F-1. Means and Standard Deviations of Teacher Responses

Institution		Cooperation Between Students and Custody Staff		Teacher's Perception of Support from Facility Administrators		Possession of State Teaching Credential or License	
	n	Mean	SD	Mean	SD	Mean	SD
1	4	3.3	.5	3.5	.6	1.0	.6
2	8	3.3	1.2	3.1	.8	1.0	.0
3	3	3.0	.0	4.0	.0	1.3	.6
4	8	2.9	.4	3.4	1.1	1.8	.5
5	6	2.4	.5	2.6	1.1	1.0	.0
6	16	3.2	.6	3.0	.6	1.0	.0
7	6	2.5	.7	3.3	.5	1.2	.4
8	-	-	-	-	-	-	-
9	3	3.3	.6	4.0	.0	1.0	.0
10	11	3.0	.5	3.1	.3	1.0	.0
11	7	3.0	.7	3.0	.6	1.0	.0
12	2	-	-	3.0	.0	1.5	.0
13	2	3.5	.7	3.5	.7	1.5	.7
14	7	3.1	.7	3.0	.6	1.0	.0
15	1	3.0	.0	4.0	.0	1.0	.0
16	8	2.8	.4	3.4	.5	1.0	.0
17	15	3.2	.8	3.4	.7	1.0	.0
18	9	2.4	1.1	3.3	.5	1.0	.0
19	8	2.3	.5	3.4	.5	1.3	.5
20	4	3.8	.5	4.0	.0	1.3	.5
21	15	2.5	.8	3.4	.5	1.1	.3
22	6	1.7	.8	2.3	.8	1.0	.0
23	3	2.3	.6	3.5	.7	1.0	.0
24	4	2.3	.5	3.3	1.0	1.0	.0
25	3	3.0	1.0	3.7	.6	1.0	.0
26	6	1.8	.4	3.2	.8	1.0	.0
27	11	1.8	.7	3.0	.4	1.0	.0
28	6	2.7	.5	3.5	.5	1.0	.0
29	2	2.5	.7	2.0	1.4	1.0	.0
30	6	2.6	.9	4.0	.0	1.0	.0

Table F-1. Means and Standard Deviations of Teacher Responses (Cont'd)

Institution		Amount of Support from Special Reading Teacher		Amount of Assistance from Counselor, Psychologist in Reading Class		Amount of Student Use of Instructional Materials in Reading	
	n	Mean	SD	Mean	SD	Mean	SD
1	4	0.0	.0	3.0	.0	34.0	.0
2	8	0.2	.4	1.8	1.0	35.0	4.2
3	3	0.0	.0	2.5	.7	33.5	3.5
4	8	0.0	.0	4.0	.0	36.5	1.7
5	6	0.0	.0	4.0	.0	33.0	.0
6	16	0.0	.0	1.0	.0	29.2	4.5
7	6	0.0	.0	2.0	.0	35.0	.0
8	-	-	-	-	-	-	-
9	3	0.0	.0	3.0	.0	35.0	.0
10	11	0.0	.0	3.0	1.4	39.5	3.5
11	7	32.0	9.8	1.8	.5	37.8	2.1
12	2	0.0	.0	4.0	.0	39.0	.0
13	2	0.0	.0	0.0	.0	-	-
14	7	0.0	.0	4.0	.0	36.5	2.1
15	1	0.0	.0	2.0	.0	20.0	.0
16	8	0.0	.0	1.0	.0	32.0	.0
17	15	0.0	.0	3.5	.7	30.5	12.0
18	9	0.0	.0	1.0	.0	32.5	2.1
19	8	0.0	.0	1.0	.0	33.7	.6
20	4	15.0	21.2	4.0	.0	26.5	2.1
21	15	0.7	.6	3.0	.0	33.0	2.6
22	6	0.0	.0	1.0	.0	-	-
23	3	0.0	.0	1.0	.0	36.0	.0
24	4	0.0	.0	4.0	.0	37.0	.0
25	3	0.0	.0	1.0	.0	36.0	.0
26	6	0.0	.0	0.0	.0	30.0	.0
27	11	15.0	21.2	1.0	.0	34.0	8.5
28	6	0.0	.0	2.0	.0	40.0	.0
29	2	4.0	5.7	1.0	.0	-	-
30	6	20.0	.0	1.0	.0	14.0	.0

Table F-1. Means and Standard Deviations of Teacher Responses (Cont'd)

Institution	n	Amount of Advance Planning in Reading		Frequency of Grading Students in Reading		Amount of Assistance from Media Specialist in Reading	
		Mean	SD	Mean	SD	Mean	SD
1	4	3.0	.0	5.0	.0	1.0	.0
2	8	2.4	.9	2.6	2.2	2.2	1.6
3	3	1.0	.0	5.5	.7	1.0	.0
4	8	1.4	.5	1.0	.0	4.0	.0
5	6	1.0	.0	1.0	.0	1.0	.0
6	16	2.0	.7	4.2	1.6	1.0	.0
7	6	1.0	.0	3.0	.0	1.0	.0
8	-	-	-	-	-	-	-
9	3	3.0	.0	5.0	.0	1.0	.0
10	11	2.0	.0	3.0	2.8	2.5	2.1
11	7	2.3	.5	3.5	2.9	4.0	.0
12	2	1.0	.0	1.0	.0	1.0	.0
13	2	-	-	-	-	-	-
14	7	1.0	.0	2.0	1.4	21.0	1.4
15	1	3.0	.0	5.0	.0	3.0	.0
16	8	2.0	.0	6.0	.0	1.0	.0
17	15	2.0	1.4	5.0	1.4	1.0	.0
18	9	1.7	.6	1.0	.0	1.0	.0
19	8	1.5	.6	6.0	.0	1.0	.0
20	4	2.0	.0	1.0	.0	-	-
21	15	1.3	.6	5.7	.5	2.0	.0
22	6	2.0	.0	5.0	.0	1.0	.0
23	3	3.0	.0	5.0	.0	1.0	.0
24	4	1.0	.0	1.0	.0	4.0	.0
25	3	3.0	.0	6.0	.0	1.0	.0
26	6	3.0	.0	1.0	.0	3.0	.0
27	11	2.0	1.4	3.5	3.5	1.0	.0
28	6	3.0	.0	1.0	.0	1.0	.0
29	2	1.5	.7	2.5	2.1	1.0	.0
30	6	3.0	.0	4.0	.0	1.0	.0

Table F-1. Means and Standard Deviations of Teacher Responses (Cont'd)

Institution		Amount of Assistance from Special Math Teacher		Amount of Assistance from Counselor, Psychologist in Math		Amount of Student Use of Instructional Materials in Math	
	n	Mean	SD	Mean	SD	Mean	SD
1	4	0.0	.0	1.0	.0	-	-
2	8	0.0	.0	1.0	.0	23.0	3.5
3	3	-	-	-	-	-	-
4	8	0.0	.0	4.0	.0	26.0	4.0
5	6	-	-	-	-	-	-
6	16	0.0	.0	1.2	.4	17.8	2.6
7	6	-	-	2.0	.0	21.0	.0
8	-	-	-	-	-	-	-
9	3	0.0	.0	3.0	.0	23.0	.0
10	11	-	-	-	-	25.0	.0
11	7	30.0	17.3	2.0	1.0	27.0	2.0
12	2	-	-	4.0	.0	23.0	.0
13	2	-	-	1.0	.0	23.0	.0
14	7	-	-	3.0	.0	25.0	.0
15	1	0.0	.0	3.0	.0	-	-
16	8	0.0	.0	1.0	.0	18.0	.0
17	15	0.0	.0	3.0	.0	13.0	.0
18	9	0.0	.0	1.0	.0	-	-
19	8	0.0	.0	1.0	.0	16.0	.0
20	4	-	-	-	-	-	-
21	15	0.0	.0	2.0	.0	23.3	2.1
22	[illegible]	-	-	1.0	.0	16.0	-
23	3	0.0	.0	1.0	.0	19.0	.0
24	4	-	-	4.0	.0	28.0	.0
25	3	-	-	-	-	-	-
26	6	-	-	-	-	24.0	.0
27	11	1.0	.0	1.0	.0	21.5	2.1
28	6	30.0	.0	2.0	.0	23.0	.0
29	2	0.0	.0	1.0	.0	20.0	5.7
30	6	9.0	.0	1.0	.0	21.0	.0

Table F-1. Means and Standard Deviations of Teacher Responses (Cont'd)

Institution		Amount of Advance Planning in Math		Frequency of Grading Students in Math		Amount of Support from Teacher Aide or Assistant in Math	
	n	Mean	SD	Mean	SD	Mean	SD
1	4	1.0	.0	5.0	.0	12.0	.0
2	8	2.5	1.0	2.0	2.0	2.7	2.5
3	3	-	-	-	-	-	-
4	8	1.9	.4	1.0	.0	6.3	12.5
5	6	-	-	-	-	-	-
6	16	2.4	.5	3.8	1.1	2.6	3.7
7	6	1.0	.0	4.5	2.1	40.0	.0
8	-	-	-	-	-	-	-
9	3	3.0	.0	5.0	.0	0.0	.0
10	11	2.0	.0	-	-	30.0	.0
11	7	2.3	.6	6.0	.0	31.7	14.4
12	2	-	-	-	-	40.0	.0
13	2	3.0	.0	1.0	.0	12.0	.0
14	7	2.0	.0	6.0	.0	30.0	.0
15	1	3.0	.0	6.0	.0	0.0	.0
16	8	1.0	.0	6.0	.0	0.0	.0
17	15	1.0	.0	6.0	.0	0.0	.0
18	9	2.0	.0	1.0	.0	8.0	.0
19	8	2.0	.0	6.0	.0	40.0	.0
20	4	-	-	-	-	-	-
21	15	1.0	.0	5.3	.6	0.3	.6
22	6	1.0	.0	4.0	.0	5.0	.0
23	3	3.0	.0	6.0	.0	0.0	.0
24	4	2.0	.0	1.0	.0	-	-
25	3	-	-	-	-	-	-
26	6	2.0	.0	4.0	.0	-	-
27	11	3.0	.0	5.5	.7	15.5	20.5
28	6	3.0	.0	1.0	.0	30.0	.0
29	2	2.0	.0	3.5	3.5	0.0	.0
30	6	3.0	.0	5.0	.0	0.0	.0

Table F-1. Means and Standard Deviations of Teacher Responses (Cont'd)

Institution	n	Amount of Student Use of Special Instructional Equipment in Math	
		Mean	SD
1	4	21.0	.0
2	8	13.5	3.5
3	3	-	-
4	8	5.0	.0
5	6	-	-
6	16	7.0	2.7
7	6	21.0	1.4
8	-	-	-
9	3	21.0	.0
10	11	-	-
11	7	10.3	6.8
12	2	28.0	.0
13	2	9.0	.0
14	7	-	-
15	1	5.0	.0
16	8	9.0	.0
17	15	12.0	.0
18	9	13.0	.0
19	8	11.0	.0
20	4	-	-
21	15	15.7	4.0
22	6	5.0	.0
23	3	5.0	.0
24	4	5.0	.0
25	3	-	-
26	6	19.0	.0
27	11	14.0	4.2
28	6	25.0	.0
29	2	10.0	.0
30	6	13.0	.0

APPENDIX G

MEAN PROPORTIONS CORRECT ON DOMAIN REFERENCED ACHIEVEMENT TESTS FOR EACH RANDOMLY SELECTED INSTITUTION

Table G-1. Arithmetic Results

Institution	Title I Plus Regular Education					Regular Education Only				
		Wave of Testing			Change from 1st to 3rd Testing		Wave of Testing			Change from 1st to 3rd Testing
	n	1st	2nd	3rd		n	1st	2nd	3rd	
1	14	.67	.65	.64	-.03	-*	-	-	-	-
2	1	.42	.42	.38	-.04	7	.61	.56	.59	-.02
5 **	1	.52	.62	.51	-.01	6	.77	.77	.73	-.04
6	14	.57	.53	.52	-.05	-	-	-	-	-
7	7	.74	.77	.73	-.01	-	-	-	-	-
8	2	.50	.46	.44	-.06	1	.72	.66	.66	-.06
9	11	.67	.65	.62	-.05	-	-	-	-	-
10	11	.50	.53	.53	.03	16	.63	.62	.64	.01
11	11	.51	.51	.53	.02	2	.78	.79	.77	-.01
12	1	.59	.54	.66	.07	1	.93	.93	.87	-.07
14 **	20	.46	.44	.48	.02	10	.60	.59	.55	-.05
15	2	.60	.65	.60	.00	2	.55	.48	.56	.01
16	5	.59	.54	.61	.02	2	.49	.79	.78	.29
17	13	.71	.72	.74	.03	8	.80	.81	.82	.02
18	2	.72	.78	.70	-.02	1	.94	.96	.86	-.08
19	8	.76	.78	.76	.00	8	.75	.82	.81	.06
20	24	.77	.75	.80	.03	8	.86	.86	.90	.04
21	17	.77	.77	.81	.04	9	.70	.71	.72	.02
22	5	.67	.66	.62	-.05	-	-	-	-	-
23	6	.68	.80	.79	.11	6	.72	.80	.79	.07
25 **	14	.64	.60	.58	-.06	5	.54	.55	.47	-.07
26	7	.66	.68	.68	.02	9	.67	.70	.69	.02
27	18	.65	.69	.69	.04	1	.85	.84	.87	.03
28	3	.77	.77	.76	-.01	4	.85	.90	.91	.06
29	2	.81	.88	.80	-.01	-	-	-	-	-
30	9	.80	.79	.78	-.02	-	-	-	-	-

* - Indicates that this institution had no regular education students with complete arithmetic data at the first, second, and third wave of testing.

** - At institutions 3, 4, 13, 24, no student took arithmetic tests at Waves 1, 2, and 3.

Table G-2. Literal Comprehension Results

Institution	Title I Plus Regular Education					Regular Education Only				
		Wave of Testing			Change from 1st to 3rd Testing		Wave of Testing			Change from 1st to 3rd Testing
	n	1st	2nd	3rd		n	1st	2nd	3rd	
1	11	.71	.64	.61	-.10	-*	-	-	-	-
3 **	1	.40	.68	.18	-.22	-	-	-	-	-
5 **	2	.71	.74	.62	-.09	7	.75	.73	.74	-.01
6	18	.56	.44	.48	-.08	-	-	-	-	-
7	10	.53	.55	.51	-.02	-	-	-	-	-
8	4	.48	.41	.40	-.08	1	.57	.44	.57	.00
9	12	.67	.64	.67	.00	-	-	-	-	-
10	16	.39	.45	.46	.07	15	.60	.67	.58	-.02
11	13	.48	.50	.51	.03	2	.92	.87	.85	-.07
12	2	.25	.19	.28	.03	1	.89	.51	.96	.07
15 **	4	.45	.30	.23	-.22	2	.80	.32	.79	-.01
16	6	.50	.44	.50	.00	4	.76	.83	.86	.10
17	13	.71	.54	.73	.02	6	.91	.86	.93	.02
18	3	.78	.81	.83	.05	2	.82	.89	.91	.09
19	7	.77	.61	.67	-.10	6	.63	.67	.55	-.08
20	29	.76	.71	.77	.01	3	.90	.86	.91	.01
21	14	.77	.66	.75	-.02	10	.72	.63	.72	.00
22	3	.48	.67	.55	.07	-	-	-	-	-
23	6	.64	.58	.62	-.02	7	.83	.86	.85	.02
25 **	12	.69	.62	.63	-.06	6	.52	.63	.58	.06
26	6	.76	.68	.75	-.01	7	.81	.70	.85	.04
27	20	.74	.80	.70	-.04	1	.79	.82	.94	.15
28	3	.77	.70	.76	-.01	5	.86	.80	.84	-.02
29	5	.66	.65	.66	.00	-	-	-	-	-
30	9	.72	.72	.69	-.03	-	-	-	-	-

* - Indicates that this institution had no regular education students with complete literal comprehension data at the first, second, and third waves of testing.

** - At institutions 2, 4, 13, 14, 24, no student took literal comprehension tests at Waves 1, 2, and 3.

APPENDIX H

IS TITLE I EFFECTIVE IN IMPROVING STUDENTS' ATTITUDES AND PROFICIENCY IN BASIC SKILLS?

RESULTS OF ANALYSES-OF-COVARIANCE

To answer this question one must be able to make the inference that observed outcomes, if any, are due to the Title I program. It is generally agreed that in order to reasonably infer a causal connection of this nature, one must conduct an experiment in which units are randomly allocated to treatments. Title I programs, however, are directed to serve the neediest of the needy; thus, true random experiments are impossible. If the process of determining which are the neediest students were known in complete detail, then a quasi-experimental design such as the regression-discontinuity model might be appropriate. However, the institutions in our sample used many different indicators of need, combining judgements, test scores and availability of space to make *ad hoc* assignments to Title I.

Cronbach, *et al*. (1976) suggest that when the complete discriminant (*all* characteristics used in assignment) is available the analysis of covariance will produce an unbiased estimate of treatment effects. They show that when neither the complete discriminant nor the complete covariate (a summary of *all* influences after assignment to treatment) is available the analysis of covariance will undercorrect or over-correct for initial differences. The direction of bias is unknown. The data collection in this study did not capture either the complete discriminant or the complete covariate. In spite of these difficulties, it is important to federal policy makers to determine whether the Title I programs have a beneficial effect. To attempt to meet this need, an analysis-of-covariance was performed, controlling for known differences in student background characteristics and initial achievement or initial attitudes between students who received Title I plus regular education instruction and those who received only regular education instruction. Inferences from this analysis must be qualified because measurement errors in the independent variables and alternative or additional causes not included in the analysis may bias the results.

All measured student characteristics were examined for possible entry into the analyses. Only students with complete arithmetic or language arts data for the first and third waves of testing were included in the analyses. In the initial analysis, achievement at the third wave of testing was compared among Title I and regular education students controlling for institution, achievement at the first wave of testing, age, sex, highest grade completed, number of schools attended, number of times institutionalized, months left in the institution, ethnic background, reason for institutionalization, and all two-way interactions among student background variables. Ethnic background and reason for institutionalization were criterion-scaled by assigning the mean achievement level at wave 3 for a category to the value of that category. For example, white students had a mean proportion correct on arithmetic of .77; the criterion-scaled value of their ethnic background then became .77.

In the initial analysis, few of the variables and interactions added significantly to predicting the achievement outcome because the initial level of achievement (wave 1 score) explained a large proportion of the variance in the outcome: 70 percent in arithmetic and 60 percent in language arts. In successive analyses only variables that had significantly contributed to explaining the variation in the outcome after initial level of achievement was accounted for were entered. Thus the final analysis included a parsimonious set of only those variables that contributed to explaining the variance in the outcome. For arithmetic the variables in the analysis included the institution, the initial level of achievement, and the sex of the student. For language arts the analysis included the institution and the initial level of achievement only.

The global analysis-of-covariance showed that, when initial differences in achievement between Title I and regular education students were accounted

for, Title I students were no different from regular education students. Thus Title I instruction seemed to have no effect, positive or negative, on achievement over the twelve week interval.*

Detailed results of the final analyses for arithmetic and language arts are displayed in Tables H-1 and H-2. The variables included accounted for the major portion of the differences between Title I and regular education students: 78 percent and 69 percent in arithmetic and language arts, respectively. It seems likely, therefore, that the important variables in explaining differences in achievement between Title I and regular education students were included in the analysis.

The results in Tables H-1 and H-2 show that arithmetic and language arts scores differed significantly across institutions. That some institutions did better than others was discussed in the fifth chapter of this report. The estimated values of the parameters in Tables H-1 and H-2 show that initial achievement (wave 1) was the most important predictor of final achievement (wave 3). Students who did well initially performed well at wave 3. In arithmetic, sex of the student was also important, female students tending to do better than males. In language arts, no significant effect of the sex of the student appeared.

The same procedures were applied to data on student attitudes toward self and school to discover differences between Title I and regular students. The four major attitude scales that were thought to relate to education outcomes were selected for this analysis: locus of control, liking of school, teacher support, and practical orientation. These scales were discussed in the third chapter of this report. A analysis-of-covariance was conducted for each attitude scale. The variables entered as predictors of wave 3 attitudes

*The analysis-of-covariance limited to the language arts scores to show a marginally significant negative effect of Title I. However, the small magnitude of this effect and the questionable validity of this analysis lead us to dismiss this result.

Table H-1. Results of the Analysis of Covariance for Arithmetic

Source of Variation	df	Sums of Squares	F Value	Significance Level	Partial Regression Estimate	Standard Error	R^2
Institution	28	7.39	16.11	.0001	*	*	
Initial Achievement	1	12.96	790.95	.0001	.834	.030	
Sex	1	0.13	8.00	.0049	-.051	.018	
Educational Program	1	0.00	0.07	.7901	.004	.013	.78

*A computational technique was used to reduce computation costs when a blocking factor with a large number of levels (institution) is a term in the model. This technique does not compute the sums of squares for that factor partialling out the effects of other variables in the model. Thus the parameters for institutions could not be estimated. The estimates computed for the other variables did take into account the effects of differences among institutions.

Table H-2. Results of the Analysis of Covariance for Language Arts

Source of Variation	df	Sums of Squares	F Value	Significance Level	Partial Regression Estimate	Standard Error	R^2
Institution	28	14.89	10.16	.0001	*	*	
Initial Achievement	1	29.17	557.23	.0001	.752	.033	
Educational Program	1	0.20	3.84	.0507	.044	.023	.69

*A computational technique was used to reduce computation costs when a blocking factor with a large number of levels (institution) is a term in the model. This technique does not compute the sums of squares for that factor partialling out the effects of other variables in the model. Thus the parameters for institutions could not be estimated. The estimates computed for the other variables did take into account the effects of differences among institutions.

were attitudes at wave 1, institution, student background characteristics and all two-way interactions among student variables. Successive analyses were conducted, each analysis eliminating nonsignificant predictors, until parimonious sets of variables were obtained. Each variable contributed significantly to predicting the outcome. Ethnic background and reason for institutionalization were criterion-scaled by assigning the mean attitude level at wave 3 for a category to the value of that category.

The results of the analysis-of-covariance are presented in Tables H-3 through H-6. The analyses showed no significant differences between Title I and regular education students on any attitude measure. Thus, to the extent that these analyses can be taken seriously, Title I instruction over twelve weeks seemed to have no effect, positive or negative, on student attitudes towards themselves and school. The results of the analyses did, however, show effects of other variables on attitudes. As was shown for achievement, differences across institutions were significant.

The variables included in these analyses accounted for less variance in wave 3 attitudes than in the analyses of achievement. In particular, initial (wave 1) and final (wave 3) attitudes were less highly related to each other than were initial and final achievement. Factors other than initial attitudes, students characteristics, and educational program seem to account for some of the variation and change in students' attitudes toward self and school.

Table H-3. Results of the Analysis of Covariance for Locus of Control

Source of Variation	df	Sums of Squares	F Value	Significance Level	Partial Regression Estimate	Standard Error	R^2
Institution	28	1041.08	3.79	.0001	*	*	
Initial Locus of Control	1	2063.04	210.03	.0001	.615	.043	
Educational Program	1	11.63	1.18	.278	.342	.314	.48

*A computational techniques was used to reduce computation costs when a blocking factor with a large number of levels (institution) is a term in the model. This technique does not compute the sums of squares for that factor partialling out the effects of other variables in the model. Thus the parameters for institutions could not be estimated. The estimates computed for the other variables did take into account the effects of differences among institutions.

Table H-4. Results of the Analysis of Covariance for Liking of School

Source of Variation	df	Sums of Squares	F Value	Significance Level	Partial Regression Estimate	Standard Error	R^2
Institution	27	251.60	2.36	.0002	*	*	
Initial Dislike for School	1	397.36	100.66	.0001	.469	.047	
Reason for Institutionalization	1	26.30	6.66	.010	.420	.162	
Educational Program	1	0.19	0.05	.827	.046	.208	.33

*A computational technique was used to reduce computation costs when a blocking factor with a large number of levels (institution) is a term in the model. This technique does not compute the sums of squares for that factor partialling out the effects of other variables in the model. Thus the parameters for institutions could not be estimated. The estimates computed for the other variables did take into account the effects of differences among institutions.

Table H-5. Results of the Analysis of Covariance for Teacher Support

Source of Variation	df	Sums of Squares	F Value	Significance Level	Partial Regression Estimate	Standard Error	R^2
Institution	28	460.91	3.89	.0001	*	*	
Initial Teacher Support	1	864.80	204.37	.0001	.587	.042	
Sex	1	37.43	8.85	.003	-.834	.281	
Educational Program	1	0.74	0.18	.676	-.088	.210	.47

*A computational technique was used to reduce computation costs when a blocking factor with a large number of levels (institution) is a term in the model. This technique does not compute the sums of squares for that factor partialling out the effects of other variables in the model. Thus the parameters for institutions could not be estimated. The estimates computed for the other variables did take into account the effects of differences among institutions.

Table H-6. Results of the Analysis of Covariance for Practical Orientation

Source of Variation	df	Sums of Squares	F Value	Significance Level	Partial Regression Estimate	Standard Error	R^2
Institution	26	184.43	4.94	.0001	*	*	
Initial Practical Orientation	1	408.87	285.02	.0001	.672	.040	
Number of Schools Attended	1	8.10	5.64	.018	-.050	.020	
Reason for Institutionalizatior	1	9.77	6.81	.010	.271	.103	
Educational Program	1	0.78	0.54	.462	.095	.130	.58

*A computational technique was used to reduce computation costs when a blocking factor with a large number of levels (institution) is a term in the model. This technique does not compute the sums of squares for that factor partialling out the effects of other variables in the model. Thus the parameters for institutions could not be estimated. The estimates computed for the other variables did take into account the effects of differences among institutions.

APPENDIX I. A COMPARISON OF THE DOMAIN REFERENCED TEST SCORES OF NEGLECTED OR DELINQUENT YOUTH WITH SCORES OF ELEVENTH GRADERS FROM THE FLORIDA STATEWIDE ASSESSMENT PROGRAM

The comparisons reported in this appendix are intended to provide further evidence about the performance of neglected or delinquent youth. The previously reported data (Chapter V) compared the standardized test scores of neglected or delinquent youth to those of grade-school students. Data on the standardized tests used in those comparisons simply are not available for non-institutionalized, high school students. In this appendix the data from the mathematics domain referenced test are compared to some of the data obtained by the Florida Statewide Assessment program for 11th grade pupils. These comparisons are not very rigorous, but they do contribute to the evidence of a considerable deficit in the performance of neglected or delinquent students.

In October of 1977 the Florida Statewide Assessment Program obtained achievement data on virtually all (about 116,000) 11th grade pupils in the state. Three of the "standards" they used in mathematics are close to some of the DRT sub-domains used in the present study.

Pupils in the 11th grade are generally one or two years younger than the average student in the neglected or delinquent institutions under study. All things other than age being equal, the 11th graders would score lower on tests than the N or D pupils. The Florida test may be easier than the DRTs used in the present study because the Florida test was multiple choice. The Florida test results indicate the percentage of pupils who passed each standard by obtaining a score of 70 percent correct or better. To obtain a score to use in comparison to the N or D test scores (expressed as the proportion correct), the Florida summaries were converted to an estimated proportion correct. This was accomplished by assuming that the group passing the standard was uniformly (evenly) distributed between the cutoff score (70%) and 100 percent correct, and that the failing group was uniformly distributed between zero percent correct and the cutoff. The estimated proportion correct may then be

estimated by using:

Estimated proportion correct =
(proportion passing the standard x assumed average score for passing group) + (proportion failing the standard x assumed average score for failing group)

An example will illustrate:

Proportion passing the standard = .72
Assumed average score for passing group is the midpoint between .70 and 1.00 = .85

Proportion failing the standard = (1-.72) = .28
Assumed average score for failing group is the midpoint between 0.00 and 0.70 = .35

.72 x .85 = .61 (The contribution of the passing group)
.28 x .35 = .10 (The contribution of the failing group)
.71 The estimated proportion correct

These estimates are probably too low because the distributions are skewed negatively, that is, there are more higher scores than lower scores. If any adjustment were to be made for this skew in the distribution, it would result in a wider gap between the N or D student scores and those of Florida's 11th graders.

Each of the three Florida standards is discussed below, including the N or D comparison data.

1. Standard H: Multiply two three-digit numbers. The Florida data indicate that the average proportion correct score on this standard wa .84. In the N or D sample the average proportion correct on all of the integer multiplication items (excluding facts--see Chapter III for more detail on the test items) was .55 among Title I students and .61 among non-Title I students.

2. Standard J (Three skills):

Add two mixed numbers.

Subtract a whole number and a mixed number.

Subtract two mixed numbers.

The Florida data provided an estimate of .69 for the average proportion correct on the combination of these three skills. In the present study, data were available on the addition and subtraction of fractions (including mixed number problems of each type, but no whole numbers). The Title I N or D students correctly answered one-third of the addition items and one-fourth of the subtraction items. The average proportions correct for non-Title I N or D students was .43 for addition and .35 for subtraction.

3. Standard M (Two skills):

Multiple two decimal numbers.

Divide decimal numbers.

The Florida data provided an estimate of .71 for the average proportion correct on the combination of these two skills. In the present study, similar items were included in the tests. The results are given separately, by skill, because there is quite a difference between multiplication and division in operations with decimal numbers.

The average proportions correct for Title I N or D students were .25 for the multiplication items and .34 for the division items. Non-Title I N or D students correctly answered an average proportion of .33 of the multiplication items and .49 of the division items. The three division items included one division of a decimal fraction by a whole number (where the placement of the decimal should be fairly easy) and one where the division of one decimal fraction into another results in a whole number (so that not placing a decimal point at all--which could be the result of not knowing anything about decimal numbers--is scored as correct).

The comparisons presented above, while lacking rigor, do indicate the very large discrepancy between the performance of institutionalized neglected or delinquent youth and their near-age peers in the regular-school population. These data are consistent with the other comparisons presented in Chapter V. The magnitude of the discrepancy in performance indicates that the N or D students will have difficulty competing with their age-peers in school and in the job market because of their deficiencies in basic skills.

The N or D results reported here are for students from the thirty randomly sampled institutions and represent scores from the first wave of testing (November and December of 1977). The Florida Statewide Assessment Program data were found in:

Stevenson, C.R. and Hoover, H.D., 1977-78 Technical Report, Statewide Assessment Program, Florida Department of Education (Undated).

A Guide to 1977-78 Statewide Assessment Results
Statewide Assessment Program, Florida Department of Education (Undated, no authorship credit).